EXPERIMENTS IN PSYCHOLOGY

WILLIAM S. FOSTER
AND
MILES A. TINKER

EXPERIMENTS IN PSYCHOLOGY

REVISED EDITION

by

WILLIAM S. FOSTER

LATE PROFESSOR OF PSYCHOLOGY IN THE UNIVERSITY OF MINNESOTA

and

MILES A. TINKER

ASSISTANT PROFESSOR OF PSYCHOLOGY IN THE UNIVERSITY OF MINNESOTA

NEW YORK

HENRY HOLT AND COMPANY

PRINTED IN THE
UNITED STATES OF AMERICA

To

EDWARD BRADFORD TITCHENER

PREFACE TO THE FIRST EDITION

An elementary course in experimental psychology is elected by a constantly increasing number of students in our universities and higher schools. This increase is apparently associated both with the increasing popular interest in psychology and with a variety of special local conditions. At the University of Minnesota, for example, such a course partially satisfies the "science requirement" for entrance to the Senior College of Science, Literature and the Arts and is recommended or advised by other colleges.

This manual grew out of our attempts to deal more adequately and at not too great expense with larger numbers of students, and at the same time to give them some appreciation both of general scientific and of specifically psychological method. It makes no pretense of completeness or finality either in form or in content, but is simply a more convenient substitute for the mimeographed sheets of instructions which we have used and gradually modified since 1919. It is published for the sake of readier availability, and in the hope that it may prove useful to others who have similar conditions to meet.

We have sought from the beginning to deal with the laboratory section as a unit. Simultaneous performance of a given experiment by all students in a section makes possible a definite sequence of the topics to be treated in the course, simplifies the work of instructors both by eliminating the necessity for constant repetition of sup-

plementary instructions and of demonstrations to individuals, makes supervision and the correction of reports more efficient, and facilitates the collection and comparative statistical treatment of data from the class as a whole. The physical arrangement of the laboratory and the choice of suitable experiments have helped in the accomplishment of this purpose.

Our main laboratory for instruction is a very large room, lighted on three sides by windows and above by a large skylight. Seats, a blackboard, and a large demonstration table occupy the center of the room and are used for lectures, introductory talks, demonstrations, and group experiments. Twelve alcoves, each containing two tables and a wide shelf beneath the window, surround this central space and open upon it. This arrangement permits the instructors and assistants to supervise individual experiments more readily than if they were conducted in separate rooms, while at the same time it sufficiently minimizes the interferences and distractions which would arise if classes of forty-eight students were to work together in an open laboratory.

The experiments described are for the most part not original, but are modifications of those commonly performed. Their selection and form were determined by the purposes mentioned, but also by the desire to avoid so far as possible the need for expensive equipment, and to emphasize, not the simple demonstration, but the methods of measurement and control which are characteristic of scientific procedure in general.

Each experiment is regarded as a project or minor problem in itself. Attempt is made to give directions in greater than usual detail, and to list precautions and warnings in more than usual completeness. Reliability is

gained wherever possible by printing a set form of instructions for the experimenter to read to the subject. Definiteness, completeness, and uniformity of laboratory reports are favored by a series of questions which call for as full and specific a treatment of data and as full an interpretation of results as seems practicable at the elementary level. A brief introduction to each experiment and a short list of strictly relevant references are intended to give the setting and to show the student the purpose of each experiment and the relationship of its results to the more general discussions in the text-books. The instructor will naturally also refer students to any particular text-books in general psychology used at his own institution.

If laboratory time is given for the collection and discussion of class results as well as for a certain amount of individual computation under the instructor's guidance, the performance of all experiments listed will probably require more than the ninety hours usually allotted. Our arrangement of experiments by topical pairs, however, allows the ready omission of some or all of the alternate experiments and the completion of an abbreviated, but a none the less fairly representative course.

I wish to acknowledge my grateful obligations to all who have contributed to the manual whether by way of suggestions in matters of experimental procedure or by more specific and direct assistance. My acknowledgments in both regards should be far more numerous than space well permits. Each one of my colleagues and former colleagues in the department at Minnesota has coöperated directly in the selection and in the trial and criticism of experiments. I am especially indebted in

these respects to Drs. Mabel R. Fernald, Mildred Loring Sylvester, Katherine Ludgate, Herbert Woodrow, Paul T. Young, and Charles Bird. To my wife and to Professor R. M. Elliott, the chairman of the department, I owe both the encouragement and the assistance which made preparation of the manual possible at this time. To them and to Professor E. G. Boring, who has read the early chapters, I am deeply indebted for specific and valuable criticism. Thanks are due to the Macmillan Company and to the author of *Experimental Psychology* for permission to quote the rules for "*How to Fail in Laboratory Work*," and to the editor of the *Psychological Review* and to the original user, Dr. Rudolph Pintner, for permission to use the photographs which accompany one of the experiments.

In dedicating the volume to Professor Titchener I seek, however inadequately, to express not only my appreciation of the training he gives his students and my sense of personal good fortune in being numbered among them, but also my gratitude for his constant aid and friendship.

W. S. F.

Minneapolis,
June 14, 1923.

PREFACE TO THE REVISED EDITION

The first edition of this book, written by the late Professor W. S. Foster, marked an important advance in the teaching of laboratory psychology. Subsequent progress in psychology and changes in methods of laboratory instruction have made a revision of the original work advisable.

Most of the statements in the preface to the first edition hold equally well for the revision. Changes in the book represent corresponding alterations in the laboratory course at Minnesota. Modifications and additions were made after experimenting for several terms with the content of the course and with methods of teaching.

It is our belief that experimental psychology should be made as interesting as possible without departing at all from sound psychological principles and scientific method. Furthermore such a course should incorporate experiments closely related to contemporary psychological trends. In this revision we have sought to emphasize the following points.

The content of the course has been increased by adding several experiments, allowing individual instructors greater latitude of choice in planning their laboratory work. The statistical computations have been reduced to a minimum by decreasing greatly the number of problems in each experiment. Mathematical tables have been included to facilitate calculations. Greater emphasis is placed on method and interpretation; less on computation. In order to eliminate unessential labor still further we strongly advise users of this book to allow

students to state in their reports that the apparatus and method are the same as in the text rather than to require that this material be copied.

The course can be made less difficult by omitting the discussion of reliability and validity in Chapter XXIII, and by omitting the whole of Chapters XXVI and XXVII. If this is done it will be necessary to leave out or modify individual problems in experiment No. 18 and in certain succeeding experiments. A few problems may be omitted from each experiment in the course to shorten the student's report.

The chapters on statistical methods are placed just after the experiments in which the methods are first used in order to emphasize the point that statistics are incidental in this course. We have found it satisfactory to avoid exercises entirely statistical and to start the course with experimentation. The statistical explanations are used merely to aid computations in the experiments.

To bring the course into closer relations with contemporary psychological trends new experiments on animal learning, human learning, special aptitude tests, and advertising have been included. In some of these experiments the use of graphic rating scales and proficiency profiles are illustrated.

Four experiments included in the first edition have been omitted and 12 new experiments added; the sections on statistics and several experiments have been thoroughly revised, and the rest of the experiments and the introduction revised to some extent.

About half of the experiments, including Nos. 1, 4, 10 and 12-24, can be used equally well in a general laboratory course and in an educational psychology laboratory course.

I wish to express here my grateful obligations to all

who have contributed to the revision by suggestions, criticisms, or direct assistance. To Professors Charles Bird, D. G. Paterson, W. T. Heron, Paul R. Farnsworth, C. P. Stone, and Mr. H. P. Longstaff I am indebted for suggestions and critical reading of parts of the manuscript, and to Professors R. M. Elliott, Josephine C. Foster and to Eva L. Tinker for a critical reading of the entire book. I am particularly indebted to the laboratory instructors at Minnesota, Dr. Mary Shirley, Dr. Kate Hevner, Mr. W. G. McAllister, and Mr. P. E. Schellenberg who have coöperated directly in the trial and criticism of experiments and in reading the manuscript. Grateful thanks are due Professor Florence L. Goodenough who read the chapters on statistics and gave specific and valuable criticism; to the Institute of Child Welfare, of the University of Minnesota, for its coöperation in securing the pictures and test scores necessary for Experiment 30. I take this opportunity to thank the users of the book for criticising the first edition. To Professor R. M. Elliott I owe both the encouragement and the assistance which made preparation of the revision possible at this time. Thanks are due Dr. Shirley for the rating scale used in Experiment 23 and to Dr. Farnsworth for the rating scale used in No. 27; to Professor C. E. Seashore for permission to use Measures of Musical Talent and to quote norms for the measures; to Professor J. P. Porter for allowing the reprint of part of *Tables to Facilitate the Computation of Coefficients of Correlation;* to Professor Joseph Peterson for permission to use his Rational Learning Test; to Mr. P. E. Schellenberg for permission to use the Schellenberg Frequency Table; and to Professor N. C. Meier for permission to use the Meier-Seashore Art Judgment Test.

June, 1929.

CONTENTS

EXPERIMENTS IN PSYCHOLOGY

CHAPTER I

INTRODUCTION

Scientific Method and Experimental Psychology

The Purpose of This Course.—This course is intended primarily to give the elementary student training in scientific method in the field of psychology. Appreciation of experimental procedure in general is not merely a matter of desirable culture for every student in this scientific age, but is more and more evidently becoming a prerequisite for investigations in such practical fields as those of business and finance, government and law, education and the social sciences, and in general wherever more reliable data, conclusions and recommendations are sought for.

The general character of experimental method is identical in all sciences, and the methodical procedures which all follow in the collection and in the interpretation of data are essentially the same in principle. In the longer established sciences, however, where elementary concepts are already a part of common speech, and where strict control of the conditions of investigation is less difficult, these very essentials may easily be taken for granted or receive but little specific attention. Psychology, however, is the youngest experimental science and it must begin with fundamentals. The complexity of the living organism, and its responsiveness to small environmental changes constantly force upon the investigator the consideration of sources of error and of the methods for

avoiding or measuring them, and constantly lead him to look for multiple causation.

Performance of the experiments outlined should serve secondly as a partial basis for more advanced work in experimental psychology, whether human, animal, differential or applied. The control of conditions necessary in any complete experiment is stressed wherever possible, but an attempt is also made to illustrate the application of a variety of special methods for the collection, treatment and interpretation of psychological data. These special methods may be adapted with little change to similar uses in allied fields. For the accomplishment of these purposes a certain minimum of training is provided in the use of graphic and statistical methods of presenting results. The course, however, emphasizes method and interpretation rather than computation.

Laboratory demonstrations may be given to great advantage with every experiment listed below. Such demonstrations are not described in detail in this manual, but the function of demonstration is recognized by the choice of as great a variety of problems and topics as the requirements of reasonably strict control and of inexpensive material equipment will allow.

Scientific Method.—Science is characterized by the circumscribed and narrowly "*existential*" point of view which it adopts towards the universe, and by the *impersonality* of its interests and purposes.

Pure science begins with the rule that only such matters as can be actually *observed* to exist shall be taken into account. It takes for granted the reality of space and time and movement and quality, and deals only with events (*phenomena*) which can be described in terms of such concepts. It disregards the questions of

values or purposes as such, and definitely refuses to deal with mystical powers or spirits. *It aims simply at the increase of our knowledge concerning events,* and not directly at the production of a more completely logical account of reality, or of more practical attitudes or more moral conduct on our part. It describes the world only from this particular existential point of view, and leaves to philosophy (metaphysics, logic, ethics, æsthetics and religion), to the so-called applied sciences or technologies, and to the practical arts, the function of furnishing us with these other sorts of guidance and information.

From the observation of phenomena, science (both applied and pure) is led by the trend of the facts to *generalizations* and to *laws,* and finally to *hypotheses* or guesses which go beyond the observed facts. When such hypotheses involve, as they often do, the supposition of things or processes not in themselves directly observable, they still possess a definite *probability* which is proportional to the amount and the consistency of the observations upon which the hypotheses are based.

This logically limited and impersonal character of science does not mean, of course, that scientific facts are practically unimportant. The history of the past four hundred years of science shows that precisely the reverse is true. It means simply that the scientist, as such, seeks the facts for their own sake; values them for their universality and adequacy in the explanation of events rather than for their practical bearings; and believes that the true interests of man are best served if scientific investigations are carried out in complete forgetfulness of practice, and in utter disregard of immediate application.

Scientific method is just as important in applied as in the pure sciences. Observations are made; generalizations and laws are formulated. There is a tendency, which is becoming more and more marked, to make practical applications of scientific findings and to do more scientific investigation for specific practical applications.

Observation and Experiment.—Observation itself consists in giving full and properly directed *attention* and in making an intelligible *report*. Scientific observation and ordinary observation are thus the same in principle. Their difference lies not in the process of observation itself, but in the *attitude or interest* of the observer, and in the direction of attention and the type of report which this attitude favors. The events of everyday life (of which our behavior and our thinking form a part) are extremely complex, and are determined by a great variety of factors or conditions about which neither the names, nor the amounts, nor the effects are known. Owing to this complexity, and to our usual extremely practical interests, we commonly overlook those very aspects of things which are of first interest to the scientist.

In his analysis of these determining factors the scientist not only has the advantage of his narrowly directed attention to guide him, but he sets out definitely to *control* and to simplify the conditions. He resorts, that is to say, to *experiment* to enable him more readily and definitely to discover the significant causal factors, and to correlate them with their specific effects.

The experimental determination of facts of experience is superior to *a priori* or opinionated judgments. *An experiment is an observation or series of observations made under conditions which are thus controlled in*

three respects. First, in order that the conditions may be simplified, the process to be investigated must be so far as possible *isolated* from known irrelevant or disturbing influences. The observation made must then be *repeated,* usually many times over, under the given conditions. This repetition is necessary in order that the influence of unknown and still uncontrolled factors (so-called accidental errors) may have many chances to balance or neutralize each other, and so that the degree of unreliability due to such errors may be measured. Finally, in the complete experiment, a *systematically varied* series of such repeated observations needs to be made. In each series a different known condition is held constant, so that in the end its specific influence upon results can be compared or measured. These three requirements of the complete experiment—isolation, repetition, and systematic variation—will meet us constantly in the succeeding work.

Scientific observation (experiment) differs, then, from ordinary observation; first, in the attitude or purpose of the observer; secondly, in the aspect or events to which attention is turned by this purpose; thirdly, in the difference in report so determined; and fourthly, in the better control of conditions under which observation is made. The difference between scientific observation and ordinary observation does not lie primarily in the fact that one observation is made in a laboratory, as the other is not, or with the use of special apparatus or instruments of precision, or with the expression of results in accurate quantitative terms. Instruments and laboratory technique and numerical measurement are important and useful, but they are secondary, and not absolutely essential features of scientific observation.

All sciences alike deal with the same subject matter, namely, the one phenomenal, observable universe which alone falls within our experience and constitutes all that we can ever *directly* know. The various sciences differ in viewing this universe from different angles, or from different *points of view,* or from different levels of observation. Science necessarily abstracts from complete reality. No single science can ever give a complete account of the world, but only an account of the world from its own particular standpoint, and the different sciences simply observe different *aspects* of the same phenomena and give different and mutually supplementary reports of what they observe.

To the onlooker, a psychological experiment is not with certainty distinguishable from a physical or a biological experiment. The differences between them are internal differences, and lie not in the externally visible performances, but in the purpose and in the direction of interest of the experimenter and observer.

The Experiment and Reports

Partnerships.—Most experiments in human psychology require two persons for their performance: the *experimenter* (E), who directs the course of experimentation, handles any necessary instruments, and records conditions and results; and the observer or *subject* (S), who is subjected to experimentation. Partnerships should be formed at the first meeting of the class, and should be held permanent as long as satisfactory work is accomplished.

Partners ordinarily serve alternately as E and as S. In short experiments this alternation may be made inside the single laboratory period. In longer experiments

one partner may serve as E for the entire experimental period, and take his place as S only in the following period or in the following experiment. In certain of the experiments (group experiments) for which sufficient apparatus is not available for individual use, or in which specially practiced technique of experimentation is required, the instructor will act as E and the whole class as S's.

The plan of this course designates the performance of a specified amount of experimental work by each member of the class in specified laboratory periods, together with the completion of formal reports thereon outside the laboratory. For this reason, and because in nearly all cases experimentation cannot begin until all subjects are present, *experiments once missed cannot ordinarily be made up,* unless both partners can work independently of the instructor at some extra period available to both. The best that can usually be done in such cases, therefore, is to require reading and study and report of references about the experiment. Tardiness, absence, and lack of coöperation, accordingly, are sins which your instructor in the psychological laboratory will find it hard to pardon.

Note-taking.—E not only keeps the records during experimentation, but he also makes later a formal report which includes these first rougher records as an appendix.

Students are required always to bring notebook and pencil to the laboratory. Odd scraps of loose paper never allow the keeping of satisfactory and loss-proof records. Neatness and accuracy are favored by the use of special cross-section paper with squares about one-fourth of an inch on a side. Special care and foresight

in note-taking at the time of experimentation will more than pay for themselves by avoiding much of the necessity for "copying over" in the formal report.

The first rule for such careful note-taking is: *label fully.* Write down at the very start the name and the number of the experiment, the date, the name of S, and notes concerning the general physical and mental condition of S, and concerning any general external conditions of experimentation which may possibly be regarded as significant. You cannot label your work too thoroughly. Diagrams, graphs, every row and column in tables of results, and all computed mathematical values especially require clear and brief designation.

The second and most important rule is: *be orderly.* If the experiment in any way permits of it, make out a tabular form at the beginning. Thereafter keep your work in place in the appropriately labeled columns or rows. Be especially careful where figures are concerned to make them legible, and to keep units under units and tens under tens, and if on cross-section paper, one digit in a square. Half of the errors in computation that are commonly made are due to neglect of this precaution. If a figure once obtained is to be referred back to later, label it and encircle it. Encircling makes it stand out on the page with special distinctness. Number the parts of the experiment and the sequence of trials as you proceed.

The third rule supports the second. It is: *don't crowd your notes.* It is a good rule to allow at least twice as much space at the margins and between the parts of the experiment as you think will be required. Something which you do not in the beginning expect is practically certain to need recording in such spaces. Make your

tables and graphs especially large, for any extensive labeling of small diagrams is likely to make them undecipherable. Save both time and space by constant abbreviation. Never write out in full such words as sensation (sn.), attention (attn.), millimeter (mm.), subject (S), experimenter (E), etc.

Formal Report.—When partners work together E alone makes a formal report, using his notes and his outside reading as a basis. In so far as time allows, S should assist E in all mathematical calculations, and should study all questions sufficiently to make sure that he also could answer them. In group experiments a formal report is required of all S's. In examinations, all students will be held for everything covered in the course whether they have written a formal report or not on the material.

This report is usually due one week after the experiment in question has been completed, and full credit is not allowed for late reports except upon approval of the instructor. This approval, if likely to prove necessary, should be obtained in advance.

The form of report is specified. It is to be made on paper eight and one-half by eleven inches ruled (for discussion) or cross-sectioned (for graphs). Wide margins should be left at both sides for the instructor's comments. Reports which are marked "not satisfactory" should be amended in accordance with criticisms, and returned for final approval.

The pages of the report should be numbered consecutively. The separate sheets should be fastened together and placed in a manila cover, with E's name, his partner's name, and the laboratory section number printed clearly on the outside.

Adhere to the following order unless otherwise instructed in the making of reports:

(1) The name and number of the experiment, the date of its performance, the names of E and of S, a statement of the physical condition of S and of significant external conditions of experimentation.

(2) *Purpose.*—Formulate in your own words, as far as possible, a clear and brief outline of the purpose of the experiment. This should include a statement of what the experiment is expected to illustrate or prove. The purpose of an experiment must be understood clearly in order to attack adequately and solve the problem involved. Be sure that you understand the relation between the different parts of the experiment and the different parts of the purpose. Never attempt an experiment until you understand its purpose.

(3) *Description of Apparatus.*—Make this description brief and in your own words.

(4) *Method.*—The method is a description of the way in which you performed the experiment. Use the first person in describing the method. It roughly answers the question, "What was done first, what was done second, third, etc.?" Your description of the method should make it possible for anyone to check the way you performed the experiment. State it briefly in your own words.

(5) *Results and Discussion.*—This includes a complete statement of the qualitative and quantitative results obtained. The quantitative results should always be put in tables where possible and expressed in graphs where this will facilitate interpretation. A complete, accurate, and intelligent discussion, explanation, and interpretation of your results in the light of your read-

ing is here demanded. The instructor will specify what questions are to be answered. The students who get most out of the experiments will be those who connect their discussion with the assigned reading and those who evaluate their results in a critical manner. Some of the questions are to be answered by reference to the readings. In the left margin of your report designate the parts of the experiment, and the exercises, or question numbers so that they stand out distinctly.

(6) *Summary and Conclusions.* In your summary include in itemized form the important points which the experimental results have brought out. In general the conclusions should answer the question: "What have I learned by performing and writing up this experiment?" The conclusion is intimately related to the purpose. Read over again the purpose of the experiment and see whether you can give a positive answer when the purpose is restated as a question. The conclusion should be restricted to statements warranted by the data at hand. One should wait until all conditions of an experiment have been varied before drawing *final* conclusions with respect to the results obtained.

(7) Your original notes should be added as an appendix, to enable the instructor to check up the accuracy of your computations and your conclusions.

General Rules for Laboratory Work.—(1) In all laboratory work, and particularly in psychological experimentation, the first essential is that *E should thoroughly understand the problem, and know the instructions, precautions, and the exact procedure to be followed before the experiment is begun.*

Neglect of this rule has spoiled hundreds of experimental series. If E does not entirely comprehend the

purpose of the experiment and the reasons for each step in the procedure, he must call upon the instructor for help. It is far better thus to *seem* to waste a few minutes at the start of the work than to risk *literally* wasting all of the work which is done during the whole laboratory period.

E is then responsible for instructing S in the details of the part S is to play, and for making sure that he understands exactly what is required of him. This may usually be done by reading to him the set formula of printed instructions. Certainty in this matter may sometimes be assured by requiring S to restate briefly the instructions in his own words.

(2) A second rule is of equal importance with the first. Having first understood them, *obey the instructions to the letter. Never in the slightest degree vary the procedure during a series of experiments in any way not definitely and specifically prescribed.* "Don't monkey with the method." If the work seems to be running not quite smoothly, it is tempting to modify the procedure in some "little way" which seems superficially insignificant. But it is just such *apparently little* changes which often turn out to have extremely great effects upon so complex an organism as the human being with whom we are experimenting. A tiny suggestion, an uncalled-for bit of fore-knowledge, the slightest possible modification of instructions may entirely alter the attitude of S, set him on the wrong track, and lead to a useless or misleading series of results, and make them incomparable with the results which have been obtained under the conditions actually specified.

It is of the very essence of experimentation that the conditions shall be so strictly controlled and the speci-

fied procedure so closely followed that another person may repeat the experiment and arrive at precisely identical results. Any variation whatever should therefore be made only with clear intention for a particular scientific, and not for a personal reason.

(3) *If results are unexpected or disappointing try to account for them; do not try to "juggle" them into regularity.* We are working in psychology with a highly complex machine, the human organism. The conditions within that machine and also the conditions of the external environment which affect it can be controlled only partially. Certain defects and certain irregularities are consequently bound to appear in spite of our best efforts to avoid or to minimize them. We assume that such irregularities, as well as the regular occurrences, have their causes. It is our duty not to ignore or to deny such irregularities, but at least to report them wherever possible, and to admit our inability precisely to ascertain their causes under certain circumstances.

Most irregularities can be avoided by sufficient attention and strict obedience to instructions and cautions. Irregularities due to our incomplete control of conditions (unavoidable slips of S's attention, disturbing stimuli and conditions in the room, and various other temporary circumstances) must be recorded as possible explanations for irregular results. Their effects must be taken into account and measured by the methods of repetition and systematic variation later to be discussed.

S should never strive to obtain a certain kind of score. A normal reaction for the situation is always desired. *In psychological experiments there is no greater honor attached to one score than to another.* But falsified scores are dishonorable. The student who strives con-

sciously to prevent a normal reaction to a situation is demonstrating a marked lack of coöperation and hence is indirectly falsifying his reports (or scores).

(4) *Always avoid suggesting to S that such and such results are what he ought to give.* To be sure, E must make certain in advance that S understands precisely what he is to do—what to attend to, how to observe, and what and how to report. But in general the less S knows in detail of the expected results the more unprejudiced and valuable his results will be. It is a commonplace that the student who already knows the answer in algebra often arrives at that answer through the very suggestion and not through correct algebraic treatment. Suggestion is far more dangerous in psychology than in algebra.

E should avoid expressing surprise or disappointment (or even marked approval) at any given result. He should remember that suggestions can be given by a tone of voice, a gesture, an expression of the face and by numerous other minor tricks of behavior as readily and as emphatically as by direct remarks. He should, therefore, proceed impersonally and rather mechanically with his conduct of the experiment, and as far as practicable postpone discussion and questions until the given series of trials are completed. If irregularities do occur in spite of all, E should seek by indirect rather than by direct questions to get S to help him to account for them.

Because of suggestion also, it is sometimes advisable not to allow certain information to come into S's hands until after the experiment. Such information is therefore printed in a special series of *Notes for Instructors.* It is further necessary in some cases to warn members of one section of the class not to discuss the nature of

the experiment or any of its results with members of a section meeting at a later period.

(5) *Strive continuously for accuracy in observation.* In the psychological laboratory data are collected by observing the activities of others or of oneself (aided sometimes by mechanical devices such as a stop watch). Difficulties arise because observation is seldom complete or entirely accurate. Securing reliable observation is an important characteristic of the scientific method. This is achieved only when the observer is consciously striving for accuracy and gives full and careful descriptions of experience and behavior. This applies equally well to all situations; whether one is describing the subjective experience of blocking when interference in habit formation occurs, tabulating times taken for trials in mirror drawing, or describing a rat's behavior in learning a maze. The laboratory student of psychology should, therefore, attempt to increase the reliability of his observations. *Accuracy is not introduced into the original observations by a refined statistical technique.*

Strict observance of the foregoing rules will overcome the chief difficulties of psychological experimentation. Neglect of them makes failure extremely easy. The rules may be summarized as follows:

1. Be orderly. Label and tabulate.
2. Understand instructions and purpose. Then begin.
3. Always follow the instructions exactly as printed.
4. Avoid giving to S any suggestions or information that may prejudice results.
5. Record results *exactly* as they occur however irregular they may seem. Try to account for them as they are.
6. Strive for accuracy in observation.

TO STUDENTS

This course in experimental psychology is first of all a course in scientific method. Now you must bear in mind at the outset that scientific method is no longer the exclusive possession of research scientists. It is the chief instrument for the control of every situation where economy of human energy is of practical importance. All sorts of miscellaneous activities and occupations from which persons get a livelihood are in process of becoming standardized and "professionalized." The core of the changes they are undergoing is the substitution of exact and quantitative measurements for guesswork and "common-sense." Leaders in business, education, medicine, politics and civics are becoming as impatient with programs or methods formulated merely from tradition or by intuition as engineers have been with magical and superstitious formulæ for controlling the physical world. Executives, personnel and factory managers, advertisers, school boards, etc., now demand carefully collected exact data upon which to base their policies.

To meet this demand some one has got to know how to *prove*. And that, in the field of psychology, is what this course teaches. Stress is here placed on controlled methods of gathering and interpreting data. There are no short cuts, and statistics cannot be avoided. On the other hand, not one experiment is included as just something to be done. Not one is without meaning for scientific method in general. Each student must decide for himself whether he will master these methods of *proof,* in the realization that opportunity will always be at hand for applying them in practical life and the professions;

or on the other hand whether science will remain for him a closed book. To know what science teaches without knowing how its conclusions can be proved is not enough. To have notions and ideas without methods for sorting those that are true and useful from those that are false and harmful is not to be educated for life in the present century.

THE EXPERIMENTS

Order of the Experiments.—This course is planned so that certain experiments furnish material to illustrate the various statistical methods used in later experiments. Frequency graphs and representative values are illustrated by computations in Experiment No. 1; learning curves in Experiment No. 13; correlation, reliability and validity in No. 18; and reliability and comparison in No. 21. When these experiments are employed it will not be necessary to use any of the statistical exercises. Each chapter on statistical method follows the experiment in which the method is first used. The individual instructor will, of course, choose experiments suited to his local situation.

Class Data.—When data from the group as a whole are desired for later treatment, have each student copy his results on an outlined blank posted on the bulletin board. For identification, give each student a number which he is to keep throughout the course.

Problems in Results.—For each experiment the instructor is to indicate which problems the students are to discuss in their reports. Reports may be shortened by omitting some of the problems.

Reduction of Statistical Treatment.—For those classes desiring to avoid some of the more difficult statistical analyses in their laboratory course we advise the omission of the discussion of reliability and validity in Chapter XXIII, and the omission of Chapters XXVI and XXVII. This would necessitate leaving out or modifying certain problems in Experiment No. 18 and in certain succeeding experiments.

CHAPTER II

Experiment No. 1

INDIVIDUAL DIFFERENCES *

When any large group of people of a given age and sex are examined for any trait, large individual differences are discovered. This is true whether the trait in question is a mental, a physical, or a personality trait. The differences discovered between the members of such a group are quantitative and the distribution of the measured or estimated values of any trait will contain very inferior, inferior, average, superior, and very superior scores or values. Careful examination of such a distribution shows that there are wide differences in scores among individuals, that the scores are continuous from one end (very inferior) of the distribution to the other (very superior), and that the individuals tend to

* The study of individual differences or variations in human capacities arose from work in the psychological laboratory soon after 1880. J. M. Cattell, working in Wundt's laboratory at Leipzig, made the first objective measurements of individual differences and was the first to suggest practical application of such results in a paper published in 1885. Apparently the term "individual differences" was first used in psychology by Cattell at this time. The work carried on in Francis Galton's Anthropometric Laboratory (founded in 1884) and Karl Pearson's Biometric Laboratory (founded in 1901) contributed greatly to the mathematical handling of data so necessary to an adequate treatment of individual differences. Binet's work with intelligence tests in France and further work by Cattell and others in America mark the firm establishment of this field of research. Interest and research in the field of individual differences have developed rapidly and produced results of great practical value for various branches of applied psychology.

cluster about a central or average ability with relatively few cases at the extremes. The general nature of individual differences can be made clear only when all members of a large group are taken into account. This is best done by representing the data in frequency tables (frequency distributions) and frequency graphs (frequency polygon and histogram). If the group is an unselected sample and *very* large the frequency polygon representing the measurements approaches in form the theoretical *curve* of distribution called the *normal curve.* With relatively small groups and where it is impossible to obtain a "random sample" the curve of distribution usually differs somewhat from the normal curve in regularity of outline and general shape. The deviation of measurable traits is best shown by the different measures of variability (AD, σ, PE, Q).

Two main groups of factors contribute to the production of individual differences: heredity and environment. Heredity or native endowment refers to that which is inherent in the germ cells from which an individual develops. Environmental factors include all those materials and influences in our surroundings which may affect human responses in any way. Broadly conceived these factors include food substances, disease, accidents, training and opportunity. In general, both heredity and environment coöperate to produce individual differences in any given trait. As yet we have no method for determining precisely the exact contribution of either of these factors.

The purpose of this experiment is to demonstrate the existence, the amounts and the nature of individual differences among college students in various performances by: (1) representing the obtained data in frequency

tables and graphs; (2) computing measures of central tendency and variability for various performances. Four kinds of ability will be measured: (a) writing H's, (b) free association, (c) controlled association (naming opposites), and (d) cancellation of digits. Individuality of color response will also be studied.

Apparatus.—Mimeographed sheets for opposites test and cancellation test (see *Notes for Instructors*).

Method.—The instructor will act as E (experimenter) and all members of the class will serve as subjects.

Part 1. *The H Test.*—Furnish each student with a piece of paper and instruct the class: "*When I say 'go' you are to begin at the top of your paper and write as many capital H's as possible until time is called. Fill the first line and then go on to the next. Use three separate strokes of the pencil for each letter as illustrated on the blackboard*" (instructor illustrates).

Be sure that the instructions are understood and then give 2 trials of 30 seconds each. Add the total number of H's written and divide by 2 for the score.

Part 2. *Free Association.*—Instruct the class: "*After the word 'ready' I shall speak a certain stimulus word. On hearing this stimulus word you are to begin at once and write as many words as you can think of until I say 'stop.' You will have a practice period of 30 seconds and then a three-minute experimental series. Write separate words, not sentences. Any words will do as 'leg, floor, rug, wool' when the stimulus word 'chair' is given. Ready for the practice trial, the stimulus word is 'table.'*" When this is done and any questions concerning the method answered, say, "*Ready for the main series. You will have three minutes this time. The stimulus word is 'home.'*"

The score is the number of words written. The results of the practice series can be used by the instructor to illustrate statistical treatment.

Part 3. *Opposites Test.*—Pass out copies of a mimeographed sheet (see *Notes for Instructors*) containing 50 words so arranged that the opposites can be written after the words. Directions at the top of the sheet read: *"When I say 'go' begin with the first word and write as quickly as possible the opposite of each word in the following list. Stop when time is called."*

Allow 1½ minutes. Read the correct responses and have each student correct his own paper. An equally good response which is different from that read is counted correct. The score is the number of correct responses.

Part 4. *Cancellation of Digits.*—Use the blanks furnished by Stoelting (No. 27011) or a mimeographed form (as given in *Notes for Instructors*). Instruct the class: *"When I say 'go' begin with the first line of digits and draw a short horizontal line through every figure 7."* (Illustrate on the blackboard). *"Do as much as possible until time is called. Work rapidly but be careful not to make errors or omit sevens. Omissions will be subtracted from your score."*

Allow 30 seconds for the trial. Using another sheet and the same procedure have the class cancel sixes. Score is the number of each digit cancelled minus the number omitted.

Part 5. *Color Naming.*—Instruct the students to write down the first single color that they think of. Then write the names of several common colors on the blackboard and obtain frequencies of these by show of hands. Add other colors to the list on the board if necessary.

Scores consist of each color written and its frequency.

Tabulate the data for each test on the blackboard, scores for men and for women separately. These data are to be copied for use in computations by individuals. Each member of the class is to write up this experiment.

Results and Discussion.

Part I. Treatment of Measurements by Graphs.

(1) Choose the most appropriate size of class intervals (for each test) and tabulate the scores of the whole class on the H test, the free association test, and the opposites test. Plot a frequency polygon to represent the material in each of the three tables. Label tables and graphs carefully.

Compare the general forms of the three frequency polygons showing in what way they are alike; in what way dissimilar.

(2) What is the nature of the normal curve of distribution? (See Gates, p. 547.) How closely do your curves (problem 1) approximate the normal curve? How do you account for the differences?

(3) Use unit class intervals and plot a frequency polygon to represent the scores of whole class on the H test. Compare this polygon with that in question (1) plotted from grouped data (for H test). Try to account for the differences found. Which curve presents the data more suitably? Why?

(4) Choose appropriate size of class intervals and prepare separate tabulations for the scores of men and of women on the free association test, using the same sized class intervals for women as for men. Compute the percentage of total measures for men and for women falling in each class interval. Then

plot percentage histograms for scores of men and of women on the same base line. State in detail what a comparison of the graphs reveals concerning the performance of men and women on this test.

(5) Represent the material obtained in the color naming test (whole class) by a bar diagram, placing the least frequently named color on the left and the others in ascending order toward the right. Describe and interpret the graph.

(6) Using the cancellation scores of the whole class tabulate the data in class intervals and then represent the cancellation of sixes in one histogram and sevens in another on the same base line. Compare the graphs in three essential respects. Why do you think you are not asked to use a percentage histogram in this problem?

Summary and Conclusions.—Summarize your results briefly and state conclusions.

Retain Part I of this experiment until Part II is completed.

Part II. Central Tendencies and Variability of Measurable Traits.

(7) Use data for whole class, and obtain the average, the median, and the mode (ungrouped data) for scores on the H test. Why do you find differences between these three measures? Under what circumstances should each be employed? Represent the average on graph of problem (3).

(8) Using grouped data (whole class) compute by the short method the average, the average deviation,

σ, and PE for scores on the H test. Plot the first three of these measures on the appropriate graph in problem (1). When should one use σ and when AD?

(9) Obtain the median for scores on the opposites test (whole class) from the ungrouped data and from the grouped data. Explain any difference discovered between the two medians.

(10) Compute Q for cancellation of sixes; for sevens. Which scores are more variable? Under what conditions would Q and PE coincide?

(11) Compare the range of men with that of women on the free association test. Of what value is this measure of variability?

(12) How much overlapping is present with scores on the cancellation of sixes and of sevens? Explain. Point out how this overlapping may be shown graphically by reference to histograms in problem (6).

(13) In the light of the above problems and by reference to Gates, pp. 543-548, state in your own words the nature of variability in measurable traits.

Summary and Conclusions. Summarize your results briefly and state conclusions.

References

A. M. Jordan: *Educational Psychology,* 1928, pp. 283-310; R. S. Ellis: *The Psychology of Individual Differences,* 1928, pp. 74-130, esp. pp. 110-130; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 530-555, esp. pp. 543-548; D. Starch: *Educational Psychology,*

1927 (Rev. Ed.), pp. 28-54; L. M. Terman: *The Intelligence of School Children*, 1919, pp. 17-91; E. L. Thorndike: *Educational Psychology*, 1914, Vol. 3, pp. 141-308; R. Pintner: *Educational Psychology*, 1929, pp. 116-122.

CHAPTER III

GRAPHIC REPRESENTATION

Repeated observations may be recorded in a series of written notes, but such records are cumbersome and difficult to grasp as a whole. Qualitative results may usually, in the end, be expressed with sufficient completeness and with far greater intelligibility by a careful *summary* which states the chief uniformities and exceptions discovered.

Qualitative statements, however, indicate simply the presence or absence, and not the degree of a given phenomenon. They are accordingly much less desirable than results given by quantitative measurement. Numerical values are not only more accurate, but they can be summarized (compared and combined) with greater definiteness and assurance. By proper statistical treatment (rearrangement, tabulation, fractionation, grouping, numerical and graphic representation, correlation, etc.), numerical and graphic values may be brought into relationships which yield generalizations that were not at all apparent from the raw data.

In order that the experimenter shall from the start be able to interpret the results he obtains, he should first gain an understanding of the elementary principles of at least two methods of representing data. In this section, therefore, we shall discuss briefly a few important forms of *graphic representation.* Later we shall take up the computation of *measures of central tendency and variability.*

Suppose that we measure, under controlled conditions, the time needed after a uniform sound for a practiced S to slip his finger from a key. Suppose that we repeat the experiment fifty times and record the fifty "reaction times" in thousandths of a second. If we were here concerned to determine more accurately the "true" length of the reaction time to sound we should, of course, perform a greater number of experiments, and use more than a single S in order to be certain that our particular conditions and particular subject were "typical" (not unusual). Suppose, however, that we confine ourselves to the above results and ask simply: "What is the reaction time to sound for our S under these conditions?"

We take the first steps towards summarization by *rearranging* the haphazardly occurring figures in order of their size, i.e., rank order. Suppose that the result of the rearrangement is the following: 106, 108, 108, 109, 111, 111, 113, 113, 113, 114, 114, 114, 115, 115, 117, 118, 118, 118, 118, 119, 119, 119, 120, 120, 120, 120, 120, 121, 121, 121, 121, 122, 122, 123, 123, 124, 124, 125, 127, 127, 127, 128, 128, 130, 130, 131, 132, 132, 135, 136.

So far we can only say that S's reaction time is not constant, but that even under the best controlled conditions it varies in 50 experiments between 136 and 106 thousandths of a second. Expressed technically, S's total *range of variation* in 50 experiments is 30 (136-106) thousandths of a second.

The figures may be expressed much more compactly and conveniently, however, in the form of a *frequency distribution.* Instead of copying each figure, we prepare a table to show the *frequency* with which each measurement occurs. Thus Table 1 summarizes the

above figures by indicating that a measurement of 115 occurs twice; a measurement of 116 zero times; 117 once; 118 four times, etc. Such a table represents the facts as completely as did the original list of figures

TABLE 1

50 Reaction Times to Sound Represented in Frequency Distribution (Unit Class Intervals)

Measure	Frequency (F)	Measure	Frequency (F)
137	0	121	4
136	1	120	5
135	1	119	3
134	0	118	4
133	0	117	1
132	2	116	0
131	1	115	2
130	2	114	3
129	0	113	3
128	2	112	0
127	3	111	2
126	0	110	0
125	1	109	1
124	2	108	2
123	2	107	0
122	2	106	1
		105	0

since the figures are merely grouped by *unit class intervals.*

The Frequency Polygon.—Graphs representing frequencies are merely frequency distributions in picture or graphic form. Their use enables one to perceive the nature of a distribution as a whole and at a glance, an obvious advantage over frequency tables. The meaning of the above distribution can be grasped more readily, then, if the tabulated data are represented in graphic

form. (See Fig. 1.) An inch or two from the bottom of a sheet of cross-section paper draw a horizontal base line. Leave about an inch between either end of the line and the vertical margin of the paper. This horizontal

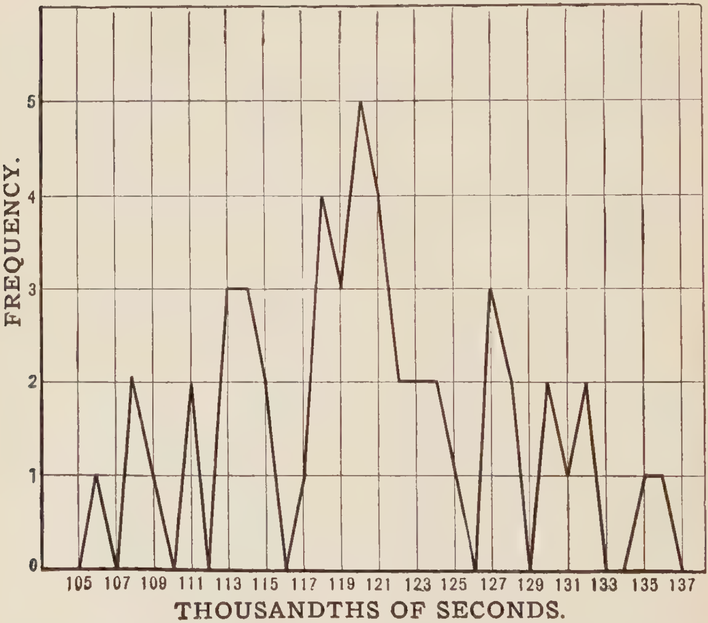


Fig. 1. Frequency polygon, representing the distribution of 50 reaction times, grouped by units as in Table 1.

line is called the abscissa scale or axis of *abscissæ*. From left to right mark off on this line equal divisions to represent the class intervals in Table 1, the lower end of the scale at the left. Use nearly the whole base line for the range of intervals. Note, however, that the first and last intervals are not located at the very ends of the base line. At the left end of this base line draw a vertical

line, the ordinate scale or axis of ordinates. Make this line into a scale by laying off on it, from below upwards, sub-division points to represent frequencies of the measures, i.e., the number of times each measure occurs (reaction times in this case). After the abscissa scale has been laid off to represent the range of measurements and the ordinate scale to represent the range of frequencies in this manner then represent the frequency, or number of times that each separate unit of measurement occurs, by placing a point at the appropriate distance above its position on the base line. Connect all of the adjacent points so determined by short straight lines. The figure thus constructed forms a *frequency polygon,* which now represents concisely and adequately all of the data originally contained in the haphazard and in the rearranged rows of figures, and later in the frequency table.

The rules for plotting the simple frequency polygon, using unit intervals, are:

(1) Note the range of measurement to be covered by the axis of abscissæ. Choose as a unit distance for this scale the greatest distance which, with the size of plotting paper available, will permit representation of all the measures as well as an additional measure beyond each end of the range. Mark off on the abscissa scale the points so determined and label them with the appropriate numbers. Note that in Fig. 1, a number labeling any point is so placed that a line projecting out from that point would *bisect* the labeling number. In this manner each number should be carefully placed *on the line* which it labels.

(2) Choose as a unit distance for the ordinate scale a distance sufficiently great to make the total graph (curve) take *relatively steep rises and dips,* and thus form a

polygon which is not too *flat*. Label the points on the ordinate scale taking the precautions mentioned in step (1).

(3) Directly above each point on the abscissa scale (*this holds only for unit class intervals*) and at distances indicated by the corresponding frequencies, locate ordinate points.

(4) Connect adjacent ordinate points by straight lines to form a polygon standing on the base line. *Note that the points representing the frequencies of the two extremes are joined to the base at the measure of zero frequency next beyond the range.* (A class interval with zero frequency is added to each extreme of a frequency distribution when a frequency polygon is to be graphed.) Work carefully that the graph may agree with the labels of the coördinates. Note in the figure that the class index "117" designates the mid-point of the interval, whose lower limit is 116.5 and whose upper limit is 117.5, and that in the polygon each point is plotted directly above the midpoint of the class interval. Label each axis to indicate exactly what it represents. Each graph as well as each table should have a number and an appropriate title.

For many purposes, however, it is not necessary or even desirable to represent exactly and directly every individual figure of a long series of experiments, either in tabulation or in the curve based upon it. For convenience and brevity we may prefer rather to *group* by class *intervals* larger than our single unit of measurement. Suppose that in the case of our reaction times we choose as class intervals: 140-144.99; 135-139.99; 130-134.99; etc. Suppose that we then determine the frequencies of the measurements in each of these classes and retabulate. The effect of such grouping is the same

as though we had recorded our original measurements, not in thousandths, but in units of five one-thousandths (or two-hundredths of a second). Table 2 represents the result of such grouping.

TABLE 2

RETABULATION OF TIMES IN TABLE 1 GROUPED IN CLASS INTERVALS OF 5

Measure	F
140-144.99	0
135-139.99	2
130-134.99	5
125-129.99	6
120-124.99	15
115-119.99	10
110-114.99	8
105-109.99	4
100-104.99	0

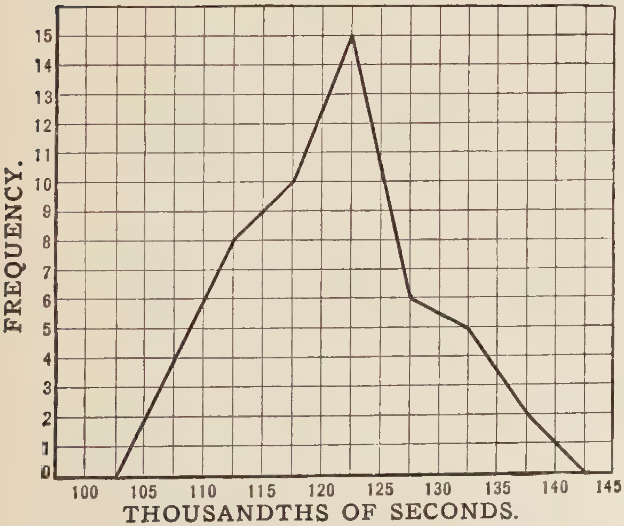


Fig. 2. Frequency polygon representing the grouped data of Table 2.

Figure 2, graphed from data in Table 2, is a more compact and regular frequency polygon than Figure 1. It is as if the irregularities of the original curve of Figure 1 had been *smoothed* by the grouping. The new curve represents the original figures in a less detailed, but for many purposes, a more convenient and significant form. If, for example, our 50 measurements were really too few to be "typical" of sound reaction times in general, or if our measurements would more sensibly have been made in two-hundredths, then our original curve was a somewhat misleading representation of a typical distribution of reaction times, for the zero frequencies at certain points in the original curve should be considered *accidental*. That is to say, if a thousand measurements had been taken instead of 50, the measures would have been found to occur *continuously* over a range of which 120 was nearly the center, and measurements of exactly 116 would probably have occurred even more often than measurements of 114 or 115. A graph representing such results would be quite *regular in outline* and similar in form to that from our grouped data. Grouping in class intervals is one method of smoothing data (the curve of distribution here).

The rules for constructing the table and the frequency polygon for figures grouped in class intervals greater than unity are the same in principle as the rules already given above. Three points, however, call for special mention:

(1) The size and limits of the class intervals should be chosen with care. Let the class intervals begin with convenient figures—0, 5, 10, 15, 20, etc., if the interval is 5; and with 0, 2, 4, 6, 8, etc., if the interval range is 2, and so on. *Note the necessity for labeling the class intervals in the distribution table exactly.* 100-104.99

means 100-104.999 + and extends *as far as* but does not include 105; the measurement 120 falls not in the class interval 115-119.99, but in the group 120-124.99, etc. Note that *intervals* are not themselves labeled on the axes (though this is permissible), but are represented simply as *distances* between adjacent labeled *points.*

In general the *size* of the class interval should be such as to result in *not less than 6 or 7 and not more than 24 class intervals in all.* Twelve or 13 is usually a desirable number of class intervals where the measures are grouped. With more numerous measures a greater number of intervals is often advisable.

(2) Haphazardly arranged figures may be ordered into groups without rearranging and rewriting them in exact order. After making out the headings of your distribution table, take such haphazard figures one by one as they stand, and for each as you go, make a check mark in the appropriate row of the table. Then count and record in the F (frequency) column the marks made in each row. See Table 4 for an illustration of this method of arranging and grouping figures.

(3) The axes of abscissæ and ordinates are scales. The figures designating the points thereon are to be placed exactly opposite the points. The class intervals are distances between such points. With group intervals greater than unity, therefore, the fair way is to *represent the frequency for each class interval by an ordinate point at the proper distance above the midpoint of the class interval,* and not above either its beginning or its end. Thus in Figure 2, the two measures between 135 and 139.99 are represented by a point two units above 137.5; the zero measures between 140 and 144.99 by bringing the curve to a point on the base line at 142.5, etc. Be very careful to follow this principle wherever

class intervals greater than unity are represented in a frequency polygon.

The Histogram.—In the frequency polygon, then, frequency is represented by a *point* (ordinate point) at a vertical distance above the midpoint of the appropriate class interval. The total number of measures in the distribution is thus represented by the sum of the ordinates. This sum is not readily taken in by eye. The frequency polygon is therefore likely to be misleading in this respect at first glance, especially if the number of cases represented is small and the range of variation is large.

Another form of graphic representation, the *histogram* or column diagram, possesses certain advantages in this, and in certain other respects which will later be emphasized. The histogram represents frequencies, not by mere distances, but by *areas* of vertical rectangles. These rectangles are erected upon the base line, using class interval distances as bases and the corresponding frequencies as ordinates. In the histogram, therefore, the number of cases is accurately represented by the *total area of the figure,* and this area can be easily comprehended by eye.

Figure 3 is a histogram which represents, like the frequency polygon of Figure 2, the grouped data of our 50 reaction times. Note that, unlike the frequency polygon, the histogram makes no use of intervals with zero frequency at either end of the distribution.

Notice particularly that in the histogram the assumption is made that the measures are not concentrated at the midpoints, but that they are *spread uniformly throughout each class interval.*

The usual rules for plotting the histogram are:

(1) Determine the size, limits, and positions of the

class intervals, and tabulate the frequency distribution exactly as in the case of the frequency polygon for grouped measures.

(2) *Between the limits* of each class interval and at a

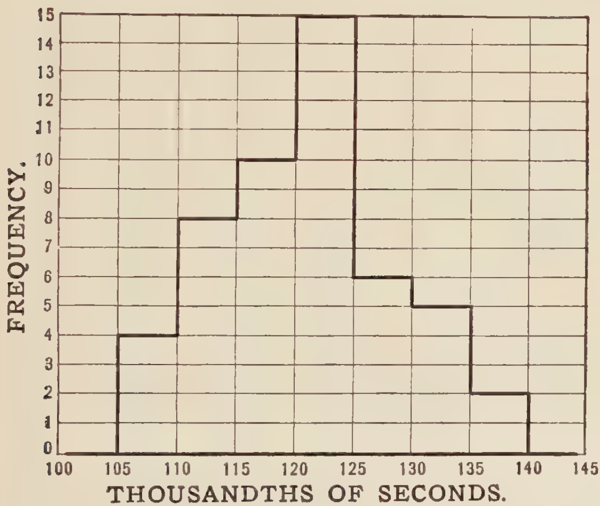


Fig. 3. Histogram representing the data of Table 2.

distance above the base line equal (on the ordinate scale) to the frequency, draw horizontal lines.

(3) Connect the adjacent ends of these short horizontals by vertical lines, and at the extremes draw verticals to the base line.

Data grouped by *unit intervals* may, of course, also be represented by a histogram. In such cases draw the rectangular columns above the midpoints on the abscissa scale, since several values all alike measured as 10 units would, if measured more accurately, have appeared scattered through a class interval extending from 9.5 to

10.499 units, i.e., the limits of the intervals whose mid-point is 10.

Military Grade	Privates	Corporals	Sergeants	Candidates for Commissions	Officers
Median Score	154	209	232	269	296

Table 3. Represents Median Scores obtained in Intelligence Test by Soldiers of Different Military Grades.

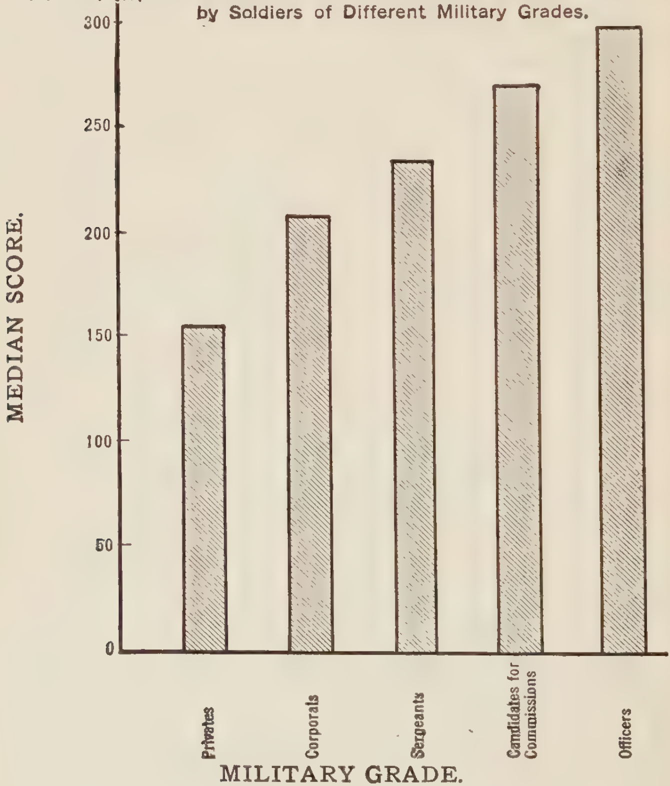


Fig. 4. Bar-diagram, representing data of Table 3.

The Bar Diagram.—In some graphs a *discontinuous* series, instead of a numerical scale, is to be represented along the base line. In such cases the use of a *bar dia-*

gram made of *separated* rectangles whose heights represent some given attribute (score, frequency, etc.), of the members of the series is theoretically preferable. An example is given in Figure 4 above. Note that there is the same amount of space between the vertical coördinate and the first rectangle on the left as there is between successive rectangles. To facilitate interpretation the rectangles should be placed in ascending or descending rank order according to their heights.

Comparative Frequency Curves and Percentage Distributions.—It is clear that a rough and easy comparison between two like distributions each involving a given number of cases can be made by plotting the two distributions upon the same scale (base line) just as if we were comparing them by placing both along a ruler.

It should be clear also that if we compare the curves for two distributions which contain *different* numbers of cases, our judgments in certain respects may easily be unwarranted, since the heights of the ordinates depend upon different numbers of cases in the two distributions. If, however, in each case we represent by rectangular heights not the actual frequencies, but the *percentages* of the total number of cases in the distribution in question, then the total areas of the two histograms will be identical (100%), and our comparisons will be fair and reliable.

Suppose, for example, that we take 50 reaction times to touch as follows: 104, 111, 78, 97, 101, 118, 100, 113, 120, 84, 107, 112, 114, 109, 132, 119, 102, 102, 90, 113, 123, 104, 113, 125, 92, 107, 99, 104, 117, 114, 105, 95, 97, 112, 109, 111, 106, 89, 105, 105, 127, 115, 108, 94, 100, 116, 121, 106, 110, 138.

Suppose that we wish to compare the above distribu-

tion with that of the 40 following reaction times to sound: 121, 110, 127, 104, 136, 127, 152, 112, 122, 132, 116, 123, 141, 115, 120, 138, 96, 145, 118, 128, 121, 121, 111, 109, 130, 122, 108, 141, 136, 134, 107, 131, 124, 119, 133, 118, 129, 125, 112, 120.

The methods of tabulation, grouping, and the percentage of times falling in each class interval are shown in Table 4. In the fourth column of the table the frequencies recorded in the third column are expressed as percentages. The comparative percentage histograms graphed from these data are in Figure 5. When two graphs are plotted on the same base line one should be drawn with a *continuous* line, and the other with a *dotted* line. The same result may be achieved by using different colored inks for the two.

TABLE 4

Shows the Method Used in Arranging Haphazard Figures, and for Grouping Figures in Class Intervals. The Figures of the Final Column in Each Part of the Table Represent the Percentages of the Total Number of Cases in Each Distribution Falling in the Various Class Intervals. The Check Marks in the Second and Fifth Column Give the Tabulation from the Original Data

	TOUCH			SOUND		
Measure	Tabulation	F	%	Tabulation	F	%
150-159.99		0	0	/	1	2.5
140-149.99		0	0	///	3	7.5
130-139.99	//	2	4	~~////~~ ///	8	20
120-129.99	~~////~~	5	10	~~////~~ ~~////~~ ////	14	35
110-119.99	~~////~~ ~~////~~ ~~////~~	15	30	~~////~~ ////	9	22.5
100-109.99	~~////~~ ~~////~~ ~~////~~ ///	18	36	////	4	10
90- 99.99	~~////~~ //	7	14	/	1	2.5
80- 89.99	//	2	4		0	0
70- 79.99	/	1	2		0	0

The Normal Curve of Distribution.—In psychological measurement, from a practical standpoint, it is impossible to measure all the human beings who might exhibit

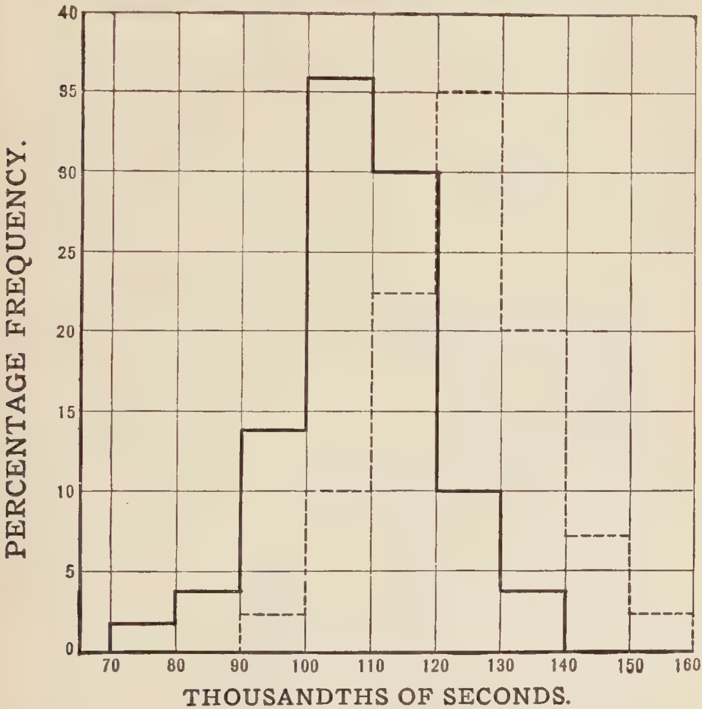


Fig. 5. Comparative histograms, representing percentage distributions of reaction times for sound and touch grouped as in Table 4.

a given trait or performance. Therefore measurements are made on a *sample* or a part of a given population. The *larger* and more *random* (unselected) such a sample is the more adequately does it represent the whole group from which it is taken. It is evident that all psychological measurements must be samplings. The more un-

selected the sample the more regular and symmetrical is the distribution curve representing the measurements.

Examination of the curves in Figures 2, 3, and 5 shows that their general shapes are somewhat similar in that a majority of the measures tend to cluster about the central area and a decreasing number extend from the crest of the curve toward either extreme. These curves are not symmetrical and have quite irregular outlines. This marked irregularity usually occurs when less than 100 cases are included in the distribution. With 300 to 400 measures the curve becomes much more symmetrical and quite smooth in outline. When several thousand measures are included the curve becomes quite symmetrical and very smooth. This leads to the conclusion that with an infinitely large number the curve would be *symmetrical* and *smooth*. Such a curve approximates the theoretical *normal curve of distribution* illustrated in

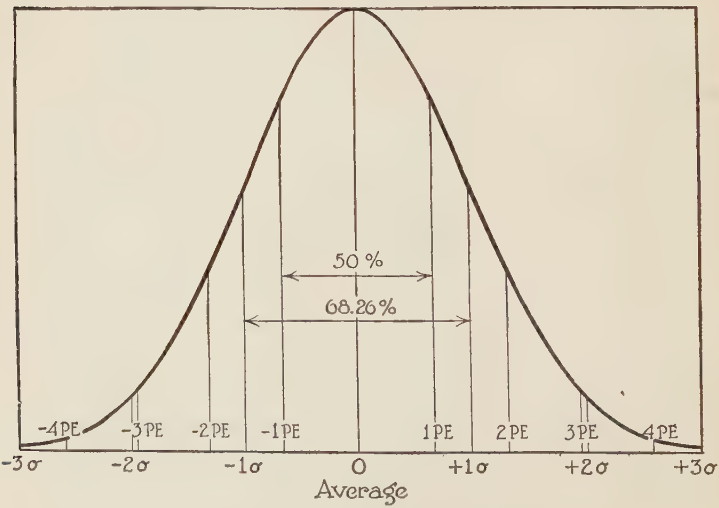


Fig. 6. The normal curve of distribution.

Figure 6. Further properties of the normal curve will be discussed later.

Precautions in Graphic Representations.—In any case where the table takes up more than one-fourth of the graph paper place the table on a separate sheet and place it in the folder *facing* the graph. Otherwise the table may ordinarily be placed at the top of the graph paper containing the frequency surface (graph). Be careful to indicate the limits of the class intervals in each row of all tables of grouped data. In such cases also remember that the ordinate points must be located above the *midpoints* of class intervals. Be sure to label *fully* and *properly* both axes, and the points upon them. Number and entitle every graph and table.

Exercises.—Use Part I of Experiment 1 to illustrate the above principles. A direct attack on the experiments without practice exercises has certain advantages. However, for the convenience of those classes which prefer practice exercises problems are given at the end of this section and other sections devoted to statistical methods.

(1) Suppose that 25 S's recall correctly the following numbers of digits in a test of immediate memory: 8, 10, 6, 6, 7, 9, 7, 7, 3, 6, 7, 5, 7, 7, 8, 6, 6, 5, 6, 8, 6, 7, 7, 9, 7. Group in unit class intervals, tabulate, and plot a frequency polygon to represent these data.

(2) In a test of range of information 30 students answered correctly the following numbers of questions: 46, 50, 27, 48, 42, 84, 48, 47, 43, 96, 39, 50, 54, 49, 40, 33, 48, 62, 41, 36, 45, 49, 51, 59, 46, 48, 37, 42, 53, 41. Rearrange the figures, group them

by class intervals of 5, tabulate, and plot the frequency polygon.

(3) The following scores were obtained in a laboratory section in a mental test: 70, 59, 46, 36, 24, 30, 34, 50, 43, 49, 51, 42, 28, 40, 39, 42, 38, 38, 28, 51, 33, 61, 41, 33, 45, 42, 44, 48, 44, 39, 18, 42, 50, 42, 66, 31, 34. Choose an appropriate class interval, tabulate, and plot a frequency polygon to represent the figures.

(4) The scores obtained in an intelligence test of 19 5-year-old boys were: 11, 11, 15, 16, 17, 17, 17, 18, 18, 19, 19, 20, 20, 22, 23, 26, 31, 31, 34. The scores received by an equal number of girls were: 11, 13, 16, 18, 18, 21, 22, 22, 23, 24, 24, 25, 25, 26, 27, 28, 28, 30, 34. Group the figures separately for boys and for girls, using class intervals of 2. On one base line, plot two frequency polygons to represent the data. Do boys or girls of this age appear to be more intelligent, if we base our judgment upon a comparison of the two curves?

(5) The following figures express in cm. the average amounts of error made by 30 students in equating the two sides of an optical illusion: 2.9, 13.4, 5.4, 3.2, 3.0, 3.8, 2.0, .7, 1.0, 7.4, 1.4, .7, .4, .5, 1.8, 9.7, 1.3, 4.0, 5.1, 2.3, .2, 8.4, .9, 3.5, .9, .5, .1, .8, .2, 6.9. Rearrange, choose an appropriate class interval, tabulate, and plot the histogram.

(6) In attempting 40 recalls of equally difficult series of jumbled letters, an S correctly recalled only 1 letter, 2 times; only 2 letters, 4 times; 3 letters, 3 times; 4 letters, 7 times; 5 letters, 6 times; 6 letters, 9 times; 7 letters, 0 times; 8 letters, 6 times; 9 letters, 2 times; 10 letters, 1 time. Fit these data

into a table, and plot a histogram to represent them.

(7) In a series of experiments on color preference it was found that Red was preferred as more pleasant on the average 12 times; Orange, 12.2 times; Yellow, 4.3 times; Green, 11.3 times; Blue, 14.5 times; Violet, 13.7 times. Represent these findings concerning the relative pleasantness of colors by a bar diagram whose abscissa represents the colors arranged as in the spectrum, and whose ordinates represent the corresponding numbers of preferences.

(8) In the next series of experiments the subjects were asked to judge concerning the unpleasantness of the same colors. Red was judged on the average more unpleasant 9.8 times; Orange, 9.4 times; Yellow, 16.3 times; Green, 9.6 times; Blue, 6.6 times; Violet, 6.3 times. Plot a bar diagram to represent the relative unpleasantness values of the colors. Compare with the bar diagram of exercise (7). Does this comparison suggest anything as to the probable relationship between pleasantness and unpleasantness?

(9) In tracing the outlines of a star seen indirectly in a mirror, 30 women required the following times in seconds: 30, 35, 43, 45, 53, 56, 60, 60, 64, 65, 70, 74, 83, 83, 97, 105, 110, 120, 120, 124, 130, 134, 140, 147, 149, 149, 170, 219, 230, 300. Twenty men required the following times: 60, 69, 80, 80, 90, 90, 92, 96, 105, 114, 115, 123, 125, 125, 160, 172, 210, 240, 254, 256. (a) Prepare separate tabulations and plot separate histograms, using class intervals of 25. (b) Compute the percentage of the total number

of women falling in each class interval. Similarly for the men. Tabulate separately. Then plot histograms for the men and for the women on the same base line, this time representing the *percentages* by the heights of the rectangles. Judging from a comparison of the percentage histograms were the men or the women slower in general?

(10) Twenty-five 10-year-old boys gave the following scores in an intelligence test: 44, 48, 52, 54, 54, 56, 58, 60, 61, 63, 65, 65, 65, 66, 66, 66, 66, 67, 67, 69, 73, 74, 76, 78, 84. Thirty 10-year-old girls gave the following scores: 39, 40, 47, 48, 50, 51, 53, 54, 55, 55, 57, 57, 58, 59, 59, 59, 59, 60, 60, 60, 66, 67, 67, 72, 72, 74, 77, 77, 79, 80. Group by class intervals of 5, compute and tabulate both direct numerical and percentage frequency distributions, but plot only percentage histograms for comparison.

References

See end of Chapter IV.

CHAPTER IV

MEASURES OF CENTRAL TENDENCY AND OF VARIABILITY

Frequency curves represent or characterize distributions of measures in a rather general way, and thus help to call attention to various aspects of them which might otherwise be overlooked. Such distributions may, however, be represented in a still more abbreviated manner, and certain specific features or tendencies of the figures may be brought out more definitely and exactly by the computation of *representative numerical values.* In this section we shall consider two sorts of representative values—*measures of central tendency* and *measures of variability.*

Measures of Central Tendency.—Repeated measurements of the same phenomenon always fall inside a certain range on the scale of measurements. The better controlled the conditions the narrower this range will usually be, and the more the measures will tend to cluster about its "center." The measures may therefore be represented by determining some such value, i.e., by computing a *measure of central tendency.*

Thus a single value may represent a whole series of measurements yielding a concise description of group performance. It also permits the "typical performance" of one group to be compared with that of others. The three measures of central tendency which we shall consider are: (1) the average,* (2) the median, and (3) the mode.

* Also called the arithmetic mean.

The Average.—The most familiar measure of central tendency is the average and may be defined as *the sum of all the measures in a series divided by the number of measures.* The formula for computing it is:

$$\text{Av.} = \frac{\Sigma \text{ M}}{\text{N}}$$

in which Av. indicates average; Σ, "the sum of," M, the individual measures; and N the number of measures in the distribution. To compute the average the individual measures need not be arranged in rank order, although, as we shall see later, other computations frequently made in connection with the average are usually facilitated by such arrangement. Computation of the average from ungrouped data is illustrated in Table 5, Case I. The measures are summed and divided by N. This procedure will be found convenient for all short series of measurements.

TABLE 5

COMPUTATION OF THE AVERAGE, THE MEDIAN, AND THE MODE FROM UNGROUPED FIGURES

Case I Scores	Case II Scores	Case III Scores
17	17	16
17	16	15
16	15	15
15	15	15
15 Med. = 15.0	15 Med. = 15.0	14 Med. = 14.5
15	14	14
14	14	13
14	13	12
13	—	—
Σ M = 136	Σ M = 119	Σ M = 114
N = 9	N = 8	N = 8

Case I. Average (Av.) $= \frac{\Sigma M}{N} = \frac{136}{9} = 15.11.$

$\frac{N}{2} = 4.5.$

Median (Med.) $=$ 15.0 for it falls on the fifth measure.

Mode $=$ Most frequent measure $=$ 15.0.

Case II. $\frac{N}{2} = 4.$

Med. $= \frac{15 + 15}{2} = 15.0$ for it falls halfway between the 4th and 5th measures.

Case III. $\frac{N}{2} = 4.$

Med. $= \frac{14 + 15}{2} = 14.5$ for it falls halfway between the 4th and 5th measures.

Where a large number of measurements are made, the easiest and most convenient way of obtaining the average is to group the measures into a frequency table and employ the "short method" for the computing. This method is illustrated in Table 6.

TABLE 6

Computation of the Averages and σ by the Short Method (Grouped Data)

(1) Scores	(2) Midpoint	(3) F	(4) D	(5) FD	(6) FD^2
135-139.99	137.5	2	+3	6	18
130-134.99	132.5	5	+2	10	20
125-129.99	127.5	6	+1	6	6
120-124.99	122.5	15	0	+22	
115-119.99	117.5	10	—1	—10	10
110-114.99	112.5	8	—2	—16	32
105-109.99	107.5	4	—3	—12	36
		N = 50		—38	$\Sigma FD^2 = 122$
				$\Sigma FD = -16$	

Guessed Average $= 122.5$.

c (correction in terms of class interval) $= \frac{-16}{50} = -.32$.

C (correction in terms of score) $= -.32 \times 5 = -1.6$.

Average $= 122.5 - 1.6 = 120.9$.

$\sigma = \sqrt{\frac{\Sigma FD^2}{N} - c^2} = \sqrt{\frac{122}{50} - .10} = \sqrt{2.34} = 1.53$ (in units of class interval).

Since a class interval consists of 5 units, then

$\sigma = 1.53 \times 5 = 7.65$ (in score units).

In this method we at first assume or guess an average which lies at about the middle of the distribution and then apply a correction to this assumed average to obtain the true average. The assumed average should be somewhere near the center of the distribution. Usually the midpoint of the interval containing the greatest frequency is in about the center and may be taken as the assumed average. Following is an outline of the steps used for the computations.

Steps used in computing the Average by the Short Method.

(1) Group the scores or measures in a frequency distribution, tabulate class intervals in column 1, their midpoints in column 2, and frequencies in column 3. Add frequencies to obtain N. N = 50 (Table 6).

(2) From inspection decide which class interval falls approximately in the center of the distribution. This is usually the place where the frequencies are most concentrated. Your estimated Av. is the midpoint of this interval. In class interval 120-124.99, midpoint 122.5 = estimated Av.

(3) Treat each class interval *as a unit* and tabulate the deviation of the midpoint of each class interval from the estimated Av., recording these deviations as positive

or negative according to whether the midpoints of the intervals are greater or less than the estimated Av. This gives column 4 (D).

(4) Multiply each deviation (D) by the frequency (F) opposite it. (Take account of signs in this and all remaining steps.) This yields column 5 (FD).

(5) Obtain the difference between the sum of the plus and the sum of the minus FD's. Algebraic Σ FD $= +22 - 38 = -16$.

(6) Divide this difference by N, the number of measures, to obtain the correction c which is in units of class intervals. $\frac{-16}{50} = -.32 =$ correction in units of class intervals (c).

(7) Multiply c by the size (number of units) of the class intervals to obtain C, the correction of the estimated Av. in scores; $-32 \times 5 = -1.6 =$ correction in score units (C).

(8) Add C (if C is negative, subtract it) to the estimated Av. to obtain the true Av.; $122.5 - 1.6 = 120.9 =$ true Av.

The obvious advantage of this "short method" of calculating the average is the small numbers involved in computations. These small numbers occur because we employ class intervals as units in figuring deviations.

The Median and Quartiles.—Another measure of central tendency which finds constant use in psychology is the *median.* It is defined as *that point on the scale of measurement upon either side of which one-half of the total number of measures will fall, if these measures are arranged in order of size* (i.e., in *rank order*). After rearrangement the computation of the median is therefore an extremely easy matter.

Finding $\frac{N}{2}$ gives the number of measures lying on either side of the median. If the number of cases is *odd*, the median is obtained simply by counting from one extreme until one finds the middlemost measure. Thus in Table 5, Case I, the median is the fifth measure, 15, for $\frac{N}{2} = 4.5$, and there are 4.5 measures either side of the center of the fifth measure. Here it is considered that the median divides the middlemost measure into two equal parts. In other words, there are just as many measures below the fifth measure as above it, namely 4.

If the total number of cases is even, no single measure in the series has an equal number of cases both above and below it, but the median is a point on the scale halfway between the two middlemost measures. Thus in Table 5, Case II, the two middlemost measures are both 15, and hence $\frac{15 + 15}{2} = 15$ is the median. In Case III, however, the two middlemost measures are 14 and 15 and hence $\frac{14 + 15}{2} = 14.5$ is the median. There are 4 measures above and 4 below the median.

In computing the median from figures grouped in class intervals we make the same assumption that we made in plotting the histogram, namely that the measures in any given class interval are distributed evenly through that interval. Look at Table 7. Notice that there are only $4 + 8 + 10 = 22$ measures which are less than 120, and that there are only $2 + 5 + 6 = 13$ measures greater than 124.99. The median (that point on the scale below which lie 25 measures and above which lie 25 measures) lies evidently somewhere on the scale between 120 and

TABLE 7

COMPUTATION OF MEDIAN AND Q FROM GROUPED DATA

Scores	F	
135-139.99	2	$\frac{N}{2} = 25.$
130-134.99	5	Median $= 120 + (\frac{3}{15} \times 5) = 121.0.$
125-129.99	6	Also Median $= 125 - (\frac{12}{15} \times 5) = 121.0.$
120-124.99	15	$\frac{N}{4} = 12.50. \qquad \frac{3N}{4} = 37.50.$
115-119.99	10	$Q_1 = 115 + (\frac{.5}{10} \times 5) = 115.25.$
110-114.99	8	$Q_3 = 125 + (\frac{.5}{6} \times 5) = 125.42.$
105-109.99	4	$Q = \frac{Q_3 - Q_1}{2} = \frac{125.42 - 115.25}{2} = 5.08.$
	$N = 50$	

The "crude" mode falls at the midpoint of the interval 120-124.99 which contains the greatest number of cases; mode = 122.5.

125, and can be found by interpolating properly between these two limits. Referring to Table 7, the steps for computing the median may be outlined as follows:

(1) Find $\frac{N}{2}$ measures, the number lying either side of the median; $\frac{N}{2} = 25$.

(2) Beginning at the end of the distribution containing small scores, count all the measures up to the interval (*critical interval*) containing the median. $4 + 8 + 10 = 22$ measures lie below the critical interval (120-124.9+).

(3) From $\frac{N}{2}$ measures subtract the sum of all measures occurring below the critical interval, $(25 - 22 = 3)$. Thus we must go the first 3 cases into the critical in-

terval in order to reach the scale point above which lie half and below which lie half of the measures.

(4) Divide the remainder obtained in step (3) by the frequency of the critical interval $(\frac{3}{15})$ and multiply the result by the number of units in the class interval. $(\frac{3}{15} \times 5 = 1.00)$. This shows that one unit on the scale within the critical interval yields the desired correction.

(5) Apply the correction by adding the final result obtained in step (4) to the lower limit of the critical interval giving the median; $120 + 1.0 = 121.0 =$ median.

Similarly, to check your results, begin at the end of the distribution containing large scores, count the number of measures lying above the critical interval, subtract this from $\frac{N}{2}$ measures to discover how many cases into the critical interval you must go to reach the median point, obtain the correction in terms of scale units and subtract this from the upper limit of the critical interval yielding the median. Calculations: $25 - 13 = 12$; $\frac{12}{15} \times 5 = 4$; $125 - 4 = 121 =$ median.

Quartiles.—The quartiles are points on the scale of measurement which also may be determined by counting or by interpolation. The *first quartile* (Q_1) is the point below which lie one-quarter and above which lie three-quarters of the measures. The third quartile (Q_3) is the point above which lie one-quarter, and below which lie three-quarters of the measures.

Inasmuch as there is considerable lack of uniformity and consequent confusion in the procedures followed in computing quartiles, it is desirable to outline one method and follow that consistently. In this book the following

procedure will be adhered to: *Always* find Q_1 by counting off 25 per cent of the measures from the end of the distribution *containing small scores* irrespective of whether small scores mean efficient or poor scores. Interpretations of Q_1 and Q_3 will depend then on the meaning of the scores; thus in intelligence test scores Q_1 falls among the poor scores. However, for time scores obtained by individuals in completing a series of arithmetical problems, Q_1 is near the efficient end of the series. The computation of quartiles is exactly the same in principle as that of the median, which is itself the second quartile (Q_2). Computation of quartiles from ungrouped data is illustrated in Table 8; from grouped data in Table 7. Further attention will be given quartiles in our discussion of quartile deviation.

The Mode.—The mode of a series is that measurement *which has the greatest frequency.* It may be determined by mere inspection of a series of scores or of a frequency table. Thus the mode for Case I, Table 5 is 15. For data in Table 6, the "crude" mode is the midpoint (122.5) of the class interval 120-124.9. The latter value appears by noting the scale measurement corresponding to the high point of the curve in Figure 2.

As above defined, the mode evidently lacks exactness, for its position is easily affected by the choice of size and limits of the class intervals, as well as by the addition of a few more "central" cases to the distribution. Certain curves, too, may show two or more modes (two or more peaks to curve). The determination of the mode, therefore, will be useful chiefly in giving us a preliminary orientation.

When to Employ the Average, the Median, and the Mode.—The *average* represents the distribution by tak-

ing into account the *actual size* of every measure. The addition to the distribution of a few very large or very small ("atypical") measures may greatly alter it, *especially when the number of cases is small.* Use the average when each score should have equal weight and when highest mathematical accuracy is sought.

The *median* can generally be obtained more quickly and easily than the average. Its position, moreover, depends mainly upon the measures of medium size and is little affected by the addition to the distribution of a few extreme measures. For distributions which contain relatively few cases, and in which a few "accidentally" large or small measures are involved, the median is to be preferred to the average as a measure of central tendency.

The chief advantage of the *mode* is its ease of determination by means of inspection of the frequency table or curve. However, because of its crudeness, it should never be used except for preliminary orientation.

Measures of Variability.—We have already indicated that averages, medians, and modes do not completely or adequately represent the frequency distributions for which they stand, but express only what their class name indicates, the central tendency. They fail entirely to represent another important feature of the series—its *variability* or dispersion. Several series of measurements, all quite different in detail, may none the less have identical measures of central tendency. Thus the average of the series 3, 3, 5, 6, 8 is $25 \div 5 = 5$, the average deviation (a measure of variability) is 1.6; of the series 1, 3, 5, 5, 7, 9 the average is $30 \div 6 = 5$, average deviation is 2.0; of the series 1, 1, 2, 5, 7, 8, 11 the average is $35 \div 7 = 5$, average deviation is 3.1. The averages

and also the medians in each series are identical. But from mere inspection (or still more clearly from graphs) the three series are seen to differ in degree of dispersion or internal variability. A computed measure of variability, the average deviation, also shows a difference for each series. Variability may be measured, and its degree expressed numerically in several ways. The most common measures are: (1) the range of variation; (2) the average deviation; (3) the standard deviation; (4) the probable error and its allied value, the quartile deviation. We shall find the greatest use for the second and third in the list.

The Range of Variation.—*The range (of variation)* of a distribution is the *distance between its least measure and its greatest measure.* It is a highly unstable and unreliable value, since the extreme measures are themselves likely to be "accidentally" large or small, and are subject to sudden changes when additional measurements are taken. Its virtue lies in the fact that it can be determined by simple inspection. It should be used only to supplement other measures of variability.

The Average Deviation.—Each individual measure in a distribution differs (varies or deviates) from the measure of central tendency by a greater or lesser amount, designated its deviation (D). The *average deviation* * (AD) is simply the *average of the individual deviations,* taken, of course, without regard to algebraic sign.

Technically there may be an AD from the average, an AD from the median, and an AD from the mode, but unless specifically stated otherwise, we understand AD to mean AD *from the average.* The formula is

* Also called the mean variation (MV).

$$AD = \frac{\Sigma D}{N}$$

In Table 8 is illustrated the computation of AD from ungrouped data. Steps in the computation may be outlined as follows:

(1) Compute the average (or median) if it has not been done. Av. = 15.3.

(2) Record the deviation (D) of each measure from this value.

(3) Add these deviations without regard to signs. $\Sigma D = 11.6$.

(4) Divide ΣD by the number of measures (N) giving $1.16 = AD$.

TABLE 8

COMPUTATION OF Q, AD, AND σ FROM UNGROUPED DATA

Scores	D	D^2
18	2.7	7.29
17	1.7	2.89
16	.7	.49
16	.7	.49
15	—.3	.09
15	—.3	.09
15	—.3	.09
14	—1.3	1.69
14	—1.3	1.69
13	—2.3	5.29

$N = 10$

$$\text{Av.} = \frac{\Sigma M}{N} = \frac{15.3}{10} = 15.3$$

$$AD = \frac{\Sigma D}{N} = \frac{11.6}{10} = 1.16$$

$\Sigma M = 153$ $\quad \Sigma D = 11.6$ $\quad \Sigma D^2 = 20.10$

$\frac{N}{4} = 2.5. \quad \frac{3N}{4} = 7.5.$

Q_1 (Falls on 3rd measure) $= 14.0.$

Q_3 (Falls on 8th measure) $= 16.0.$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{16.0 - 14.0}{2} = 1.00$$

$$\sigma = \sqrt{\frac{\Sigma D^2}{N}} = \sqrt{\frac{20.1}{10}} = \sqrt{2.01} = 1.42.$$

The Standard Deviation.—Another measure of deviation is the standard deviation, designated by σ. It is a measure of deviation from the central tendency just as the AD but it is always derived from the average and never from the median or mode. In computing σ the difficulty of neglecting algebraic signs which occurs with AD is avoided by squaring the separate deviations. We may define σ as the square root of the average of the squares of the deviations taken from the average of the distribution. For ungrouped measures σ is obtained by the formula

$$\sigma = \sqrt{\frac{\Sigma D^2}{N}}$$

Computation of σ from ungrouped data is illustrated in Table 8. Use table of squares and square roots in Appendix to facilitate computations. The steps to be taken are:

(1) Record the deviations from the Av. as in computing the AD.
(2) Record the squares of the deviations (D) in a column headed D^2.
(3) Sum these values giving $20.1 = \Sigma D^2$.
(4) Divide ΣD^2 by the number of measures (N) and then extract the square root of the quotient giving 1.42 the standard deviation (σ).

For grouped data the standard deviation is most conveniently found by the "short method." This computation is illustrated in Table 6. Steps in the computation may be outlined as follows:

(1) Tabulate frequencies, record deviations from the estimated average and obtain c (correction in units of class intervals) as in computing the true aver-

age. (See Table 6 above; $-16 \div 50 = -.32 = c$.) Square this giving c^2. $(-.32)^2 = .10 = c^2$.

(2) Multiply each FD by D to obtain FD^2 values. Add these to get ΣFD^2 which equals 122.

(3) Divide ΣFD^2 by N and subtract c^2 from the quotient. Then extract the square root of the value thus obtained. This gives σ in units of class intervals;

$$\sqrt{\frac{122}{50} - .10} = 1.53.$$

(4) To turn the σ thus computed into score units, multiply by the size of the class intervals; $1.53 \times 5 = 7.65 = \sigma$ in score units.

For short series and when the long method of computation is used the AD can be calculated somewhat more rapidly than σ. But where the "short method" of computing the average is employed there is only a little additional work required to obtain σ. The standard deviation (or the probable error which is derived from σ) is ordinarily used in research where great accuracy is desired. For, as a rule, σ is more accurate than the AD. This is because the "Standard deviation is, in general, less affected by chance fluctuations than the AD, and is, therefore, a more stable measure of dispersion." * The standard deviation should be used when reliability of discovered differences are later to be calculated for σ is employed for this. Ordinarily the use of the standard deviation is preferred to that of the average deviation.

The Probable Error.—The probable error (PE) of a distribution is another measure of variability. It is obtained directly from σ. The formula is

* Garrett, p. 27.

$$PE = .6745\ \sigma$$

Since σ must always be computed first there is no logical reason for employing PE, but its wide use in scientific literature demands a knowledge of its meaning. When 1 PE is measured off above and below the average on a normal curve the central area set off by these limits incloses 50 per cent of the measures in the distribution. The PE is put to practically the same uses as σ. This will be discussed further under "properties of the normal curve."

The Quartile Deviation.—The quartile deviation or semi-inter-quartile range (Q) *is half of the distance on the scale of measurement between the first and third quartile.* The formula is

$$Q = \frac{Q_3 - Q_1}{2}$$

Where we can assume a normal curve Q is identical with PE of the distribution and marks off the limits of 25 per cent of the scores just below and 25 per cent just above the median. PE and Q are both less than the AD which in turn is less than σ. This fact supplies a rough check on the accuracy of computed measures of variability.

As compared with the total range of variation, Q is little affected by the addition of extreme measures to the series, and is consequently a much more stable and reliable measure. It possesses the great advantage over other stable measures that its computation is easy and does not involve the necessity of first obtaining the individual variations. Table 8 illustrates the computation of Q from ungrouped data. The steps follow:

(1) Find $\frac{N}{4}$ and $\frac{3N}{4}$. If $\frac{N}{4}$ and $\frac{3N}{4}$ yield whole numbers as 2 and 6, the quartiles fall between scale values, i.e., halfway between the second and third measure and halfway between the sixth and seventh measure. But if $\frac{N}{4}$ and $\frac{3N}{4}$ yield fractions as 2.5 and 7.5 the quartiles fall on scale values, i.e., on the third and on the eighth measure. In Table 8, $\frac{N}{4} = 2.5$, $\frac{3N}{4} = 7.5$.

(2) Start at the end containing small scores and count in on the distribution to the point in the series below which lie ¼ and above which lie ¾ of the measures. This is Q_1 which $= 14.0$.

(3) Start at the same end (containing small scores) and count in on the distribution to the point in the series below which lie ¾ and above which lie ¼ of the measures. This is Q_3 which $= 16.0$.

(4) Subtract Q_1 from Q_3 and divide by 2, yielding Q, the quartile deviation; $Q = \frac{16 - 14}{2} = 1.00$.

The steps employed to compute Q from grouped data are illustrated in Table 7 and may be stated as follows: (Q_1 and Q_3 are computed by the same method as the median which is Q_2.)

(1) Find $\frac{N}{4}$ measures. $\frac{N}{4} = 12.5$.

(2) Beginning at the small end (low scores) of the distribution count all the measures up to the interval (*critical interval*) containing Q. Critical interval is 115-119.9.

(3) From $\frac{N}{4}$ measures subtract the sum of all measures

occurring below the critical interval (12.5 — 12 = .5). Divide this remainder by the frequency of the critical interval and multiply the result by the number of units in the class interval; $\frac{.5}{10} \times 5 = .25$.

(4) Add the final result obtained in step 3 to the lower limit of the critical interval, giving Q_1; $115 + .25 = 115.25 = Q_1$.

(5) In a similar manner compute Q_3 after finding $\frac{3\,N}{4}$ measures; $125 + (\frac{.5}{6} \times 5) = 125.42$.

(6) Subtract Q_1 from Q_3 and divide by 2 giving the quartile deviation (Q); $\frac{125.42 - 115.25}{2} = 5.08 = Q$.

The smaller any given measure of deviation (σ, PE, AD or Q) the more the scores are concentrated around the central tendency and hence the less the variability present. And the less the variability the greater the reliability of central tendency, for the central tendency is then more representative of the group as we shall see later.

Properties of the Normal Curve.—Refer back to Figure 6. The normal curve is bilaterally symmetrical. The line (average) drawn perpendicular to the base from the highest point of the curve divides the total area into two parts equal in area and similar in form. Perpendiculars erected at $+1\,\sigma$ and $-1\,\sigma$ distance from the average cut the curve on either side at the point where it is changing from convexity to concavity. The area included between these perpendiculars ($\pm 1\,\sigma$) contains approximately the middle ⅔ (68.26%) of the

cases. Also the area included in $\pm 3\,\sigma$ contains practically all cases (99.73%) in the distribution.

In the same way the PE which is derived from σ, is laid off to the right and left of the average. The area between perpendiculars at ± 1 PE distance from the average contains the middle 50 per cent of cases. In the normal curve PE coincides with Q. The area within the range ± 4 PE comprises 99.3 per cent of the cases.

Similarly the area between perpendiculars at ± 1 AD from the average (not shown in figure) contains the middle 57.5 per cent of cases. All these measures of variability are assumed to hold where a group of data approximates the normal distribution.

In normal distributions certain constant relations are present among measures of variability. The more important of these are:

$$\mathrm{PE} = .6745\,\sigma \qquad \mathrm{AD} = .7979\,\sigma$$
$$\mathrm{PE} = .8453\,\mathrm{AD} \qquad \sigma = 1.2533\,\mathrm{AD}$$

Graphical Representation.—Measures of central tendency and measures of variability may be represented graphically on the distribution plots by solid or dotted lines, erected vertically through the curve from the appropriate points on the base line.

Thus in Figure 7 are represented the average and σ for the data in Table 6. After the perpendicular representing the average (120.9) has been erected at the appropriate place on the base line, measure off $1\,\sigma$ distance (7.65) on either side of this and erect perpendiculars representing $\pm 1\,\sigma$. In a similar manner medians, modes, AD's and PE's are graphed. In the normal curve the average, the median and the mode all coincide.

Overlapping of Distributions.*—An important supplement to measures of central tendency and variability in group comparisons is the amount of overlapping present in the distributions of two groups on the same test or trait. Measurement of overlapping is usually done by calculating the per cent of scores in one group which

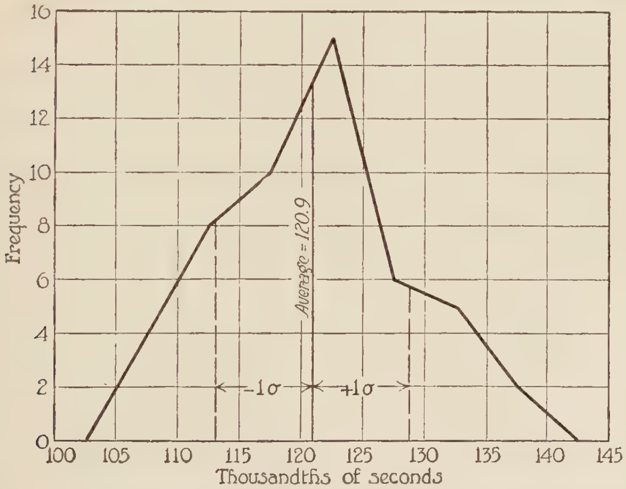


Fig. 7. Graphic representation of measures of central tendency and measures of variability. Average = 120.9; $\sigma = 7.65$.

reaches or exceeds the median of the other. If the two distributions are identical, 50 per cent of the cases in each will reach or exceed the common median. If they are not identical, more than 50 per cent of the cases in the distribution with the greater median will reach or exceed the lesser median in value, and the greater the difference in medians, the greater the excess, if other things remain equal. Overlapping is easily shown graph-

* Included in this section because it will be used in the early experiments.

ically by plotting the two frequency polygons or histograms on the same base line and drawing in the medians. (Instructor will illustrate.) To calculate overlapping proceed as follows: Obtain the medians of the two series of measures. Then, in the group with the larger median, add all scores which fall below the median of the other group. (It is not always necessary to use the group with the higher median as a base. Sometimes, because one group is of primary interest, it is desirable to reverse the process. In such a case one would obtain for the group with the smaller median the per cent of cases which reach or exceed the larger median. This would yield a percentage less than 50.) Subtract this sum from N yielding the number of measures which reaches or exceeds the smaller median. Express this figure as percentage.

Suppose, for example, that the median score of 50 boys on an arithmetic test is 48; the median score for 50 girls is 44. Suppose further that 20 scores of the boys (group with the larger median) fall below the median of the girls. Then N — 20 or 50 — 20 = 30 cases of the boys reach or exceed the median of the girls. Turning this into per cent we have 60 per cent of the boys' scores reaching or exceeding the median of the girls (the smaller median). This indicates a large amount of overlapping for the two distributions since 60 is only 10 points from 50 per cent of overlapping which occurs when the two curves of distribution are identical (and we assume that the distribution is normal). Thus the greater the difference between the per cent of overlapping and 50 per cent, the less the overlapping.

Exercises.—In the exercises below, refer back to the data and tables of your previous exercises.

Keep your computations correct to *one* decimal point. Thus, call 5.13 = 5.1; take 5.75, or 5.78, or 5.84 as 5.8, etc.

Remember that the computation of representative values is facilitated, and errors in computation are reduced by arranging the figures in rank order at the start. Computation is also facilitated by use of the table in Appendix to obtain squares and square roots.

(11) Compute the average and the median in exercise (2), Chapter III. Compare. Indicate and label their positions on the graph for that exercise.

(12) Compute the average, the median and the mode in exercise (1). Compare the three values.

(13) What is the median in exercise (3), obtained directly? Find the median from the use of the grouped figures in that exercise, and compare the two values.

(14) Using grouped data of exercise (4), compute Av. and σ for boys, for girls. Do girls or boys of this age show greater variability on intelligence test scores? What is PE of boy's scores? Of girl's scores?

(15) Obtain the Q in exercise (2), computing the quartiles from the rearranged figures direct. Again compute Q, this time obtaining the quartiles from the grouped figures. Compare the two values and explain.

(16) Compute the average, and the σ for boys in exercise (10). Compute the corresponding values for girls. Judging from these results, which appear more intelligent at the age of ten—boys or girls? Which appear to vary more in intelligence?

(17) Compute the per cent of overlapping between scores of men and of women in exercise (9). Explain.

(18) Compute average by the short method for grouped data in exercise (5).

(19) Compute AD and σ for data of exercise (1). Compare and explain why they are different.

References

H. E. Garrett: *Statistics in Psychology and Education,* 1926; A. S. Otis: *Statistical Method in Educational Measurement,* 1925; H. O. Rugg: *Statistical Methods Applied to Education,* 1917; L. L. Thurstone: *The Fundamentals of Statistics,* 1925.

CHAPTER V

Experiment No. 2

THE KNEE JERK

By analogy with the reflection of light by a mirror, the term reflex has been applied to those cases of action in which the neural impulse, initiated at a sense organ, appears to be reflected directly back from the nerve centers to an effector (muscle or gland), and there to produce a quite definite and unlearned response. A reflex action (often called simply a "reflex") is, then, a relatively *simple, definite, unlearned* and *involuntary* action.

Most behaviorists regard the reflex as the elementary unit of neural function, and hold that the more complicated forms of behavior can be understood as complexes or as modifications of reflexes. Thus instincts, attitudes, "sets," desires, and the like are regarded as complex coördinations of reflex tendencies (or impulsions), and habits and voluntary actions are regarded as modified or conditioned reflexes, or combinations of reflexes. From this point of view the young animal begins life simply as a bundle of partially independent reflex tendencies, which by modifications of neural conduction paths become better and better coördinated, so as finally to produce the more intelligent and self-sufficient behavior of the adult.

On the side of consciousness the reflex may be and may remain an entirely unconscious matter, as in the case in the pupillary reflex and in glandular secretions, but

stimulus and movement may both be consciously represented, as in the knee jerk. With conflict or interaction between reflexes or impulsive tendencies in individual development, consciousness may become clearer, as in cases where voluntary control of previously involuntary movements arises. However, the "degeneration" of conscious and voluntary action into so-called secondary reflexes such as balancing and walking, which are typically unconscious or only dimly conscious, is of at least equally frequent occurrence. We cannot therefore by any means say that the reflex is an unconscious action. It is merely a simple involuntary or uncontrolled action.

In theory the essential structures concerned in a reflex in a higher organism are: (1) a sensory cell or ending; (2) an afferent (sensory) fiber; (3) an efferent (motor) fiber; and (4) an effector cell (in muscle or gland). In the living human being the "reflexes" available for study possess no such simplicity. Even in the case of the decerebrate and the spinal frog and in the relatively simple nerve-muscle preparation where a single muscle and its nerve are dissected out, we fall far short of attaining it. Stimulation commonly affects many sensory endings; the impulses thus aroused run over nerves consisting of many fibers, through centers composed of a multitude of synapses or of association fibers which make numerous and variable connections with other centers, thence back through cables of motor fibers, and finally affect a large number of effector cells in muscle or gland.

Yet more important in practice is the fact that in their course over this "reflex arc" the primary impulses are subjected to a variety of influences. Other simultaneous nerve impulses may inhibit (check) or facilitate (rein-

force) them through combination at the nerve centers. Varying chemical and electrical conditions throughout the whole arc (sense organ, nerve and center, effector) are also known to be effective in modifying reflexes. Such conditioning factors bear the names of adaptation, set, practice, tonus, fatigue, voluntary control, and the like. Most muscles, too, are arranged in opposing pairs (flexor and extensor), and are mechanically or functionally connected with other nearby structures. The visible response which we commonly call a reflex action is, therefore, the resultant of the interplay of a truly enormous number of constantly varying factors.

Since these internal variable factors are difficult or impossible to control and since many external factors are, although to a somewhat less degree, it will not be strange if in the following experiment the disinterested, impatient, and careless experimenter and subject fail entirely to secure consistent results. Indeed, a very good reason for attempting the experiment at this time is that it demonstrates so well the necessity for serious and painstaking work.

While it has been noted that increasing the strength of a stimulus applied to a nerve increases the amount of muscle contraction, it is fairly well established that the steps of increase in contraction correspond to the number of nerve fibers supplying the muscle. This indicates that the degree of contraction is dependent on the number of nerve fibers in a state of excitement at one time, not on the strength of the impulse traveling over a single fiber. Experiments have been performed which seem to indicate that if a nerve fiber is excited at all it is excited to its maximal capacity. This is known as the *all or none principle* of nervous conduction. Therefore,

when a muscle is being excited to different degrees of activity by blows or shocks of different intensities applied to a nerve, it signifies that the stronger blows are exciting more nerve fibers (and consequently more muscle fibers) to activity, not that a stronger impulse is transmitted along the nerve.

Human reflexes include: *Muscle shortening* responses (knee jerk, etc.); *sensory adaptive* responses (contraction of pupil, focusing of eyes, etc.); *protective* responses (winking, coughing, sneezing, starting, wincing, etc.); *internal muscular* responses (swallowing, peristalsis, heart-beat, breathing, etc.); *secretory* responses (mouth-watering, dryness of mouth, "cold sweat," tears, etc.); other *emotional* responses (facial expressions, and others too numerous to mention). More complete lists are given in the references below, but all classifications are somewhat arbitrary, incomplete and uncertain.

Because of its accessibility and ease of measurement, the patellar reflex (involuntary knee jerk) has been chosen for study in this experiment. It can be easily elicited in about 98.5 per cent of normal persons, and its variations in connection with other symptoms have diagnostic importance in certain nervous disorders. Its activity is, in many respects at least, typical of reflexes in general.

The amount of the patellar reflex has been shown to vary with a great variety of factors; the nature of *stimulation* (strength, place, frequency, etc., of blow), the *position and muscular attitude* of the subject; and a great variety of *external and internal factors* affecting the higher centers and the nervous system as a whole (health, fatigue, hunger and meals, weather and time of day, attention, expectation, emotion, pain and other

simultaneous sensory excitations, voluntary actions, etc., etc.). Individual differences are great and variations due to unknown causes ("accidental errors") are common and large in amount, but certain generalizations can be made from individual results, and still better ones from the results of the class as a whole.

The purpose of this experiment is to make a quantitative determination of the extent of the knee jerk and how it changes when the following conditions are varied: (1) strength of the blow delivered on the patellar tendon, (2) the muscular attitude of the subject (tensing muscles in part of the body).

Apparatus.—Reflex hammer. Table clamp, two ordinary right angle rod clamps, one universal clamp. Adjustable support holding pivoted stick. Scale upon which movements of the upper end of the stick may be measured. Shoe attachment. Rubber bands. (See Appendix for detailed description.)

Method.—Have an instructor decide by trial which partner has the more active and hence more easily measurable knee jerk. This partner becomes S, whether or not he makes the formal report.

Seat S squarely on the end of a table so that the hinges of the knees are fairly close to the edge, and the legs swing freely. Fasten the shoe attachment to his left shoe and ankle with about four rubber bands. One tin strip is to be fastened just in front of the shoe-heel, and the other is to project out horizontally behind the lower end of the stick.

Let S move close to the scale support so that his left leg, hanging loosely, is as close as possible to the pivoted stick, yet not quite touching it. Now move the stick towards you. Have S swing his legs and allow them to

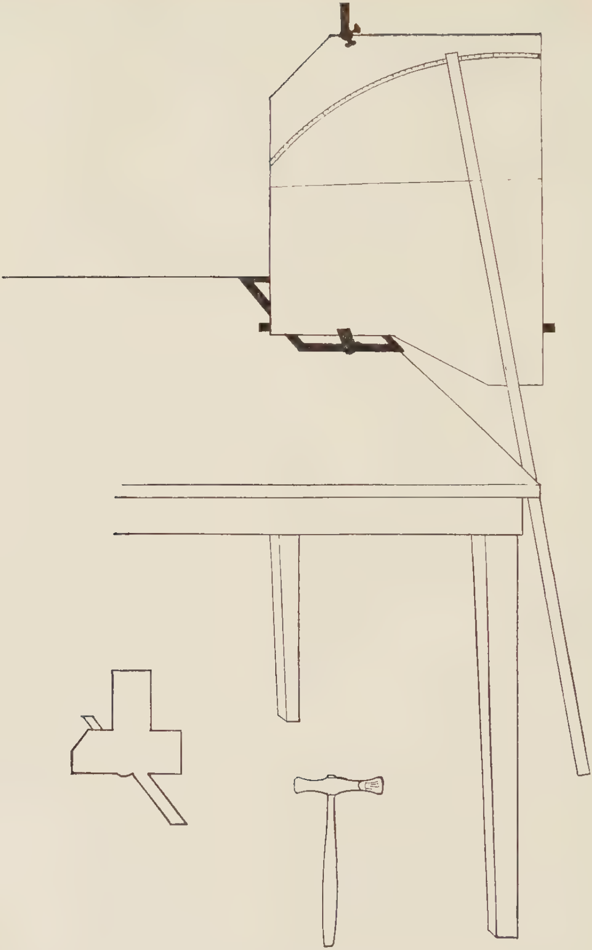


Fig. 8. Apparatus for measuring extent of knee jerk.

"die down" to a loose and natural position. Return the stick to a zero reading. Then release the table clamp and move the scale support to a position such that the projecting arm of the shoe attachment just touches the stick, still reading zero.

Instruct S as follows: *"Close your eyes. Pay as little attention and give as little thought to my preparations and actions as you can. If properly given, the blows are not painful in the least and should cause no nervousness. Relax as much as you can and try not even to anticipate or tense yourself in preparation for the stimulus, but simply let your leg take care of itself. Change your position as little as comfort permits during a series of trials."*

Feel the position of S's knee-cap through the clothing and place the tips of two fingers of your left hand on its lower "corners." The appropriate place for the blow is on the middle of the tendon immediately below the fingers. Blows too high or too low will strike bone, and such blows if repeated may rapidly become unpleasant. When the proper place for striking the tendon has been discovered, place a chalk or pencil mark on the clothing over that spot to guide the blows. The proper blow is *decided and definite, but not hard.* It is best given with the hammer held rather loosely near the end of the handle, and is rather *"lively"* and *"bouncing,"* not a "dead" blow.

A few trials should teach E to strike properly and fairly uniformly. Let ten seconds or so elapse between blows, so that fatigue of the reflex may be at least partially avoided. If S leaves his position on the table between series, have him *return to his original* (standard) *position and attitude* as exactly as possible. Changed

position or attitude may materially affect the amount of reflex, and may thus obscure or counteract the effects of the particular conditions which we wish to measure and compare. Do not let S obtain knowledge of his results until the eightieth trial is completed.

Prepare and label eight 2-column, 10-row tabular forms as follows, heading each column with a letter (N, H, I or T):

1-10 N	41-50 T
11-20 H	51-60 I
21-30 I	61-70 H
31-40 T	71-80 N

N indicates *normal* knee jerk according to instructions already given.

H indicates that E is to give noticeably *harder blows,* other conditions remaining the same as in the N series.

I indicates that in each trial before striking E is to warn S, and S is to *try voluntarily to hold his leg still* each time. The subject should attempt to inhibit the jerk, not by saying to himself, "I will not let my leg jerk this time," but rather by contracting the antagonistic muscle which is in the calf of the leg. (Contracting the muscle in the calf raises the heel and lowers the toe.) E is to give blows of normal strength only.

T indicates that before each trial S is to get ready to *tense* his arm. A second or less before striking, E is to call out "tense!" and S is to tense his left arm strongly, doubling up his fist, contracting the muscles of the upper arm and clenching his teeth in the effort. E uses blows of normal strength only.

First, take 20 preliminary trials of the "normal" reflex, recording the measurements to the nearest degree

of the scale in a table labeled "Preliminary." This practice is to accustom S to the experimental conditions and to stabilize his attitude. It is also necessary in order to practice E in the giving of proper and uniform stimulation. Obtain the median and the AD from the median. If the AD is not greater than 1/4 of the median, proceed with the series of trials outlined above. If the AD is greater than 1/4, show your results to the instructor who will inform you whether to take more practice trials before going on with the experimental series. Fair uniformity *must be acquired* in practice trials before satisfactory results can be obtained in the experimental series.

Notice the "ABBA" (really ABCDDCBA) plan upon which the above sequence of series is arranged. If we should take 20 trials with condition N, 20 with H, 20 with I and 20 with T, we should not be sure whether the discovered differences are due solely to the differences in nature of stimulus and attitude of S, to the practice of S in adjusting to experimental situation and E in delivering more uniform blows, or partly to one and partly to the other factor.

If, however, we make 10 trials with N, 10 with H, 10 with I, 10 with T, then 10 more with T, 10 with I, 10 with H and finally 10 with N again, as the ABBA schema requires, the 4 sets are at least approximately "equated" for practice, and any discovered difference can be attributed with some assurance to the difference in the experimental conditions.

This schema assumes, of course, that the effects of practice in a given number of trials are even throughout the whole sequence. Such may not be exactly the case, and sometimes it may be better to fractionate the 80 trials further and take 5 trials each in the following

order: N H I T T I H N, T I H N N H I T. The degree of fractionation necessary or desirable depends on how easily S can shift from one to another condition without "disturbance," and on how much greater are the effects of practice in early than in later trials.

Results and Discussion

(1) Obtain the median separately for the first and for the last N series (trials 1-10 and 71-80). Compare the two medians and state whether or not you regard practice and fatigue as highly significant in their influence upon the amount of normal knee jerk.

(2) Tabulate all 20 N measures; all 20 H measures; all 20 I; and all T measures. (Group measures by class intervals of 2 degrees if this results in not over 10 classes, otherwise choose larger class intervals.)

Draw a base line the full length of your cross-section paper. On this base line plot in pencil, pen, dotted, and dash lines (or in different colored inks) 4 frequency polygons to represent the distributions of the N, H, I, and T measures. Label the graph clearly and fully. Express verbally what your graphs show.

(3) Obtain the medians of the 20 N, 20 H, 20 I, and 20 T measures (ungrouped data). Compare these 4 medians, and state how much the H, I, or T conditions appear to *reinforce* or to *inhibit* the normal knee jerk. Which condition (H, I, or T) appears to give the greatest difference from normal?

(4) Measurements of the knee jerk under our conditions are likely to be very variable, and comparative

measurements rather uncertain. Cite all the controls used in this experiment and why they were employed. Suggest all the improvements you can think of in the technique, arrangements and extent of series, etc., which you think might improve the certainty of results in such an experiment. What do you consider the *chief* sources of error (reasons for uncontrolled variability) in our measurements?

(5) What *per cent* of the 20 H measures *reach or exceed* the median of N? What *per cent* of the I measures *are less* than the median of N? What *per cent* of the T measures *reach or exceed* the median of N? On this basis, what condition appears to bring the least overlapping with the N measures? Compare these results with results in (3) above. How might the results in (3) and (5) be combined in a graph to show overlapping?

(6) In terms of *nerve impulses and neural connections* involved, try to suggest why the knee jerk varies: (1) with intensity of blow; and (2) with simultaneous muscular tensions as clenching teeth and tensing arm. (Ask instructor for help on this if necessary.)

(7) Compare your subject's records with the following results obtained from 70 experiments, pointing out similarities and trying to explain marked differences:

Median N value (approx.) ....17°
Median H value (approx.) ....24°
Median I value (approx.) ....12°
Median T value (approx.) ....20°

% H reaching or exceeding median N . . 91
% I less than median N 69
% T reaching or exceeding median N . . 77

Summary and Conclusions. Briefly summarize your results and state conclusions.

References

On the relations of reflexes to behavior in general, see R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 230-233; 521-550; C. J. Herrick: *Introduction to Neurology*, 1916, esp. pp. 11-36; E. B. Holt: *The Freudian Wish*, 1915, esp. pp. 47-99. For a typical classification, see H. C. Warren: *Human Psychology*, 1919, pp. 95-111; J. F. Dashiell: *Fundamentals of Objective Psychology*, 1928, pp. 154-170.

For the neural structures and connections concerned in reflex action, see Herrick, op. cit., esp. pp. 56-182, or J. B. Watson: *Psychology from the Standpoint of a Behaviorist*, 1919, esp. pp. 113-159.

On the nature, coördination, inhibition and reinforcement of nerve impulses, see C. S. Sherrington: *The Integrative Action of the Nervous System*, 1911; K. Lucas: *The Conduction of the Nerve Impulse*, 1917; W. M. Bayliss: *Principles of General Physiology*, 1924, (4th Ed.), pp. 376-435.

On the knee jerk (technique and conditions), see C. H. Howell: *Text Book of Physiology*, 1920, pp. 140-160; and W. P. Lombard, *American Journal of Psychology*, Vol. I, 1888, pp. 5-71.

CHAPTER VI

Experiment No. 3

THE SIMPLE REACTION *

The reaction experiment, as commonly performed, requires that a subject shall respond quickly to a stimulus by making a specified movement.

Interest in such experiments began in the earliest days of experimental psychology. It arose from an astronomer's discovery of a supposed "personal equation," a difference in the speed of reaction of observers in determining the instant of passage of a star across the cross-hair of a telescope, and early investigation centered on the question of personal reaction-times.

It was shortly discovered, however, that a very great variety of factors influence both the nature and the speed of reaction, and that personal differences, although they exist, are of far less importance than are other *conditions of experimentation.* From the point of view of the psychologist, therefore, reaction-times are not so important in themselves, but because they are convenient indicators or measures of the influence of such conditions; and the reaction experiment is itself less important for its revelation of individual differences than for the bearing of its results upon more general questions of the psychology of action.

Reaction times and the reaction consciousness vary:

* Part 3 may be omitted at the discretion of the instructor to shorten the experiment. If this is done, modify questions to correspond to Part 1 and 2 only. Note that the last paragraph under Part 3 in method applies also to Part 1 and 2.

(1) with the *quality of stimulus* (light, sound, touch, printed or spoken words, etc.); (2) with its *duration*, its *intensity*, its *size*, etc.; (3) with the *problem* set the subject before reaction (simple "sensing," "discriminating" between two or more stimuli, reading and understanding a word, recognizing a given character, etc.); (4) with the individual subject and with his age, training, interest, attention, practice, fatigue, and other such *"personal" and "subjective" factors;* (5) with the *character of the response required* (pressing a key, "selecting" a proper key to press, responding by an associated word of a particular sort, remembering a learned response, etc., etc.).

Both depend, that is to say, primarily, upon the nature of the *stimulus* and upon the nature of the *"set"* (attitude, task, purpose, determining tendency) of the subject. It is because it throws so clear a light upon "set" that the reaction experiment is here chosen for performance.

The *"simple" reaction,* which involves a simple sensory stimulus, relatively simple instructions, and a simple finger movement, is so called to distinguish it from the more complex forms of reaction which involve "discrimination" among stimuli, or "association" or "choice" or the like on the side of response. It is evidently a far more highly developed form of behavior than the "true" reflex or than the so-called reflexes such as the knee jerk, breathing or heart beat.

(1) It is a *learned* and not a *native* response, and involves not only lower, but also higher nerve centers.

(2) Special *preparation on the subject's part* and special instructions and *"ready"* signals are necessary if reliable and uniform results are to be obtained.

(3) Roughly speaking, it is a "*voluntary*" and not an "*involuntary*" response. "Voluntary" and "involuntary," however, are vague and scientifically misleading words. In the concrete they can be applied only to extreme and clearcut cases of action. After practice, moreover, the subject's response is apt to become "habitual" or "automatic" and the subject certainly no longer consciously "decides" or "wills" to act in each case. He finds, rather, that his finger "reacts of itself," and that he presses the key much as in daily life he starts at a sudden loud and unexpected sound. The reaction may thus temporarily become what is sometimes called a "secondary" (or "conditioned") reflex.

The purpose of this experiment is to study voluntary action in the "simple reaction" experiment by comparing the different reaction times obtained under the three following conditions of set: (1) sensorial reaction, (2) muscular reaction, and (3) reaction with indefinite preparation.

Apparatus.—Vernier chronoscope (for description of a simple model, see Appendix). Stop watch or other timer. Ruler or stick for striking key.

Method.—The instructor will explain the principle and the adjustment of the vernier chronoscope, and may demonstrate other forms of the vernier and other chronoscopes.

E and S then adjust the chronoscope as follows:

(1) Test the longer (nearer) pendulum to see if it makes *precisely* 75 *complete* (*back-and-forth*) *swings per minute,* i.e., one swing in $^{60}/_{75}$ or $^{40}/_{50}$ of a second. Release the pendulum as the timer hand is passing zero, count "one" when the bob completes its first full swing, and continue counting for one minute.

If necessary, lengthen the pendulum slightly as follows: (a) Loosen the thread in the bob; (b) Bend the bob sidewise so that one side is lower than the other; (c) Unwind the shorter thread in the bobbin at the top end, and level the bob; (d) Rewind the loosened end about its bobbin. If it is necessary to shorten the pendulum, proceed in a similar way.

(2) See that the other (shorter) pendulum makes *approximately* 77 swings per minute, i.e., approximately one swing in $^{39}/_{50}$ of a second.

(3) Now, by releasing the two keys simultaneously, start the two pendulums swinging together, and count the swings of the longer until a second coincidence occurs. The pendulums, as they swing back and forth, will gradually get "out of step," and then gradually approach coincidence again. Judging by eye, *apparent* coincidence will usually occur on three successive beats—actual coincidence occurs on the *middle* beat of the three. It is safest, therefore, to keep counting until divergence is about to occur, and then to record the count which fell on the middle beat of the three apparently coincident swings.

If the pendulums have been adjusted for $^{40}/_{50}$ and $^{39}/_{50}$ seconds exactly, it is clear that coincidence should occur in the 39th beat of the long pendulum. Since the shorter pendulum was adjusted only approximately, however, our count will probably not be precisely 39. We must, therefore:

(4) Shorten (or lengthen, if necessary) the shorter pendulum very carefully until coincidence occurs precisely on the 39th beat of the longer.

If now E starts the long pendulum by striking its key, and S reacts to the sound so made by pressing the

other key, E needs only to count the swings of the longer from release to coincidence of the two pendulums in

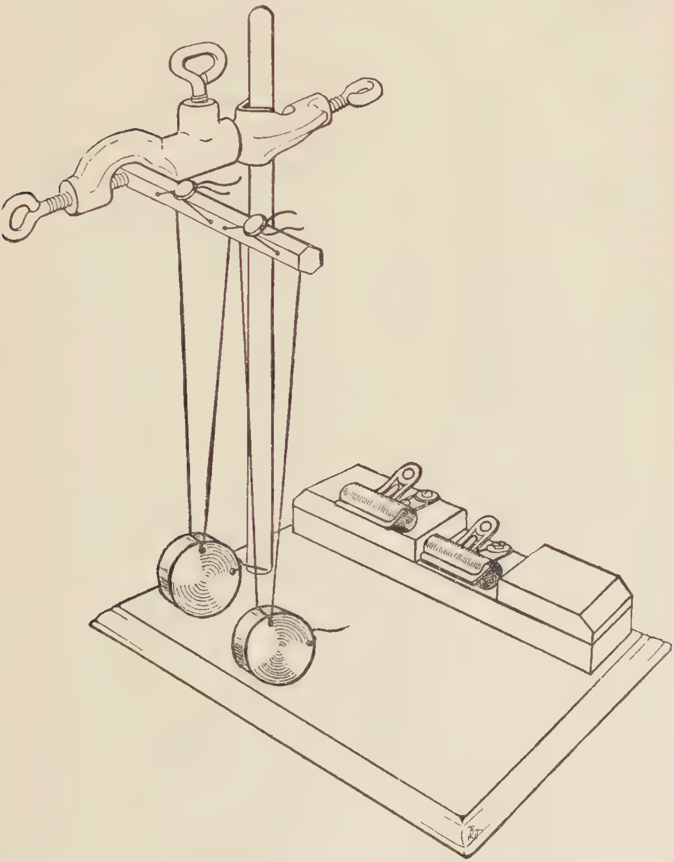


Fig. 9. Simple reaction pendulum.

order to determine S's reaction time in fiftieths of a second. For if the shorter pendulum is released very shortly after the longer one (say $7/50$ of a second), it

will catch up at the rate of $\frac{1}{50}$ ($\frac{40}{50} - \frac{39}{50}$) second per swing, and in this case catch up precisely on the seventh swing.

Seat S at the table with the forefinger of his right (or naturally used) hand resting upon the key of the *shorter* pendulum. Sit at right angles to him so that you can strike the other key conveniently and so that you can most readily note coincidence of the threads.

Get from an instructor directions as to a tabular form for recording times, introspective notes, etc., directions as to use of false stimuli, etc., and have him test your ability to give fairly uniform signals—"Ready"—pause 1½ sec.—"Now"—pause 1½ sec.—Strike key.

Part 1. *The "Sensorial" Reaction.*—Read S the following instructions for the so-called "complete" or "*sensorial" reaction,* and make sure that he understands them:

"*I shall say 'Ready!', 'Now!' and then make a sound by striking the key of the long pendulum. Close your eyes and put your attention upon listening for that particular sound. As soon as you hear that sound clearly, react by pressing the key. Do not pay attention to the movement—that will take care of itself well enough.*

"*In order to see if you are properly set and properly attentive, I shall sometimes strike a book instead of the key. You should not react to such a sound. If you do, you are probably not obeying the instructions to listen for the sound.*

"*So remember, put your attention on listening for the expected sound, and when you hear it, react by pressing the key.*"

Hold the ruler almost vertically and strike a blow

which is decided enough to sound clearly and to release the pendulum definitely. Do not strike so hard as to bend the lever of the key. Remember to count at the *end* of the first and at the end of the later *full* swings, to continue counting *through* coincidence, and to record the *middle* beat of the three approximately coincident ones.

Take preliminary practice trials until you are sure that you know how to handle the apparatus and record the counts accurately, and until you are reasonably certain that *S will not respond to "false stimuli."*

Then take about 60 trials (50 regular, and about 10 irregularly arranged false stimuli). Record the 50 *times, any remarks or behavior of S that will throw light on unusual results, and make notes as to whether S did* or *did not* react to the *false* stimuli, etc.

If (as this regular series progresses) S reacts to *any* false stimulus, remind him again of instructions. Note by an arrow in your table an interruption of the series, and practice S again until he feels sure he will not so react. Then continue the regular series again. If four or more such interruptions occur the series is to be considered *not reliable.* In such case discontinue it and take an entirely new series of 60 trials.

Part 2. *The Muscular Reaction.*—Read S the following instructions for the "abbreviated" or "muscular" reaction and make sure that he understands them:

"In this series, close your eyes as usual, but put your attention on the intended finger movement. As soon as the sound comes, react as quickly as you can by pressing with that finger. Do not pay attention for the sound—hearing will take care of itself well enough.

"When false stimuli come, you may very likely react to them unintentionally, since your attention is not upon them but upon the movement.

"Remember, direct your attention toward the intended movement, and as soon as the sound comes, react as quickly as possible."

Give S preliminary practice. Note that if a reaction time *appears* to be $\frac{4}{50}$ of a second or less, S is probably pressing his key *prematurely,* just *before* the sound comes to him. Indeed, with uniform ready signals, some S's may react at a given time after "Now," even if no sound stimulus occurs at all. If your S shows evidence of giving premature reactions, caution him against them.

After the preliminary practice, take a regular series of 60 muscular reactions including 10 to false stimuli. Note and repeat premature reactions, if any occur. Note introspections, times, remarks, etc., as in part 1 above.*

Part 3. *Indefinite Preparation.*—Get from the instructor a tabular form for arranging and recording 75 trials; of which 25 are to be with ½ sec., 25 with 1½ sec., and 25 with 5 sec., preparation intervals. These intervals are to be used in *haphazard order* so that S shall be only *indefinitely prepared* for reaction.

Instruct S, *"I shall say 'Ready' and 'Now' as usual*

* In order to equalize practice and in order to make possible an exact comparison of sensorial and muscular reaction times, 50 more trials with the muscular and then 50 more trials with the sensorial instructions should be taken, completing an ABBA arrangement. The instructor may require this addition if time permits.

The reason why we have not taken 4 sub-series of 25 trials each (instead of 2 with 50 trials each) is that the *shift* from the muscular back to the sensorial set is very difficult and uncertain for many unpracticed subjects, and the resulting errors may well be of greater importance than the error of practice involved in the 50 S—50 M arrangement suggested.

(about 1½ seconds apart), but after 'Now' I may give the stimulus quickly, moderately soon, or quite late. Put your attention upon the finger movement and when the sound comes, react as quickly as possible. Seventy-five such muscular reaction times with a haphazard order of preparation intervals will be taken."

Use a timer to control your "ready" signals. Give "Ready" as the hand is 1½ seconds from a division mark; "Now" as it crosses the mark; and "stimulus" at the designated time (½ sec., 1½ sec. or 5 sec.) thereafter. Be sure that S gets no hints (secondary criteria) as to which of the three intervals is to be used. Watch your tone of voice in particular for a "give-away."

Have S describe any differences he has discovered between his "set" for the sensorial and muscular reactions, and between definite and indefinite "set." In particular have him describe and distinguish his muscular attitudes (eye and facial muscles, breathing, finger, hand and arm muscles, etc.), in the different situations.

Results and Discussion.

(1) Summarize S's introspections (verbal reports) and his generalizations as to any differences he has discovered between his "set" for the *sensorial* and for the *muscular reaction*, and between muscular attitudes under the different sets.

(2) Compute your subject's *average sensorial* reaction time. The typical value for highly practiced subjects is about .23 sec. Convert your average, which is expressed in fiftieths of seconds, and compare your subject's average with the typical time. Try to explain any decided disagreement.

Compute the *average muscular* reaction time for

S. Compare with a typical value of .12 sec. and explain if necessary.

How much *difference* does your S show between average sensorial and average muscular time? What reasons can you suggest for this difference? Explain *in detail*, referring to instructions and other conditions of experiment and to introspections.

(3) Compare the variability of sensorial reaction (50 reactions), muscular reaction, and reaction with indefinite set (75 reactions). Explain differences.

(4) Compute the percentage of overlapping of sensorial and muscular reaction times. Show the overlapping graphically and explain what this result tends to show.

(5) Compute average reaction time for the 25 cases of indefinite preparation with $\frac{1}{2}$ sec. interval, with $1\frac{1}{2}$ sec. interval, with 5 sec. interval. Compare and explain. Compare the average for $1\frac{1}{2}$ sec. under indefinite set with average for muscular reaction and explain.

Compute average of the three averages ($\frac{1}{2}$, $1\frac{1}{2}$, 5 sec.) to find average reaction time with indefinite preparation for your S. Compare with muscular reaction time and explain.

(6) What rules would you make concerning signals for one who wishes properly to "start" the participants in a 100-yard dash? Can you suggest why a starter should use a pistol rather than a flag signal or a word of mouth?

(7) By reference to concrete features of the reaction experiment, illustrate: (a) a *stimulus* which initiates a set; (b) *preparatory responses* involving the

sense organ, neural arcs and muscles chiefly concerned in an end reaction; (c) differences between the *conscious set* of an untrained and of a practiced subject; (d) the persistence of set; (e) *facilitation* and *inhibition* of neural impulses in set; (What is the natural or reflex response to a sound? What response would not occur unless S was set? What stimuli which might attract attention and cause response are neglected if set is strong? What do premature reactions show about the facilitation of certain impulses? etc.); (f) *efficiency* and *speed* of response as related to definiteness of set.

(8) Discuss the history of the reaction experiment, its various forms, and the experimental facts concerning the dependence of reaction time upon stimulus and upon set. (See Titchener reference below.)

(9) What *physiological* explanations are given for the initiation and persistence of set? (See Woodworth and Dunlap references below.)

Summary and Conclusions.

References

For the history, technique, and facts of the reaction experiment, see E. B. Titchener: *Text Book of Psychology*, 1911, pp. 428-447; *Experimental Psychology*, 1901, Vol. I, Pt. 1, pp. 117-125; Vol. I, Pt. 2, pp. 212-227; 1906, Vol. II, Pt. 1, pp. 141-195; Vol. II, Pt. 2, pp. 326-392; C. S. Myers: *Text Book of Experimental Psychology*, 1911, pp. 125-135.

For the general bearings of the reaction experiment, see E. B. Titchener: *Text Book of Psychology, loc. cit.*; R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 234-

240; H. C. Warren: *Elements of Human Psychology,* 1922, pp. 307-311; F. H. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 3-4, 45-48; G. T. Ladd and R. S. Woodworth: *Physiological Psychology,* 1911, pp. 470-499.

For physiological theories of set, see Woodworth, *op. cit.,* esp. pp. 237-240, or K. Dunlap: *Journal of Psychobiology,* Vol. 2, 1920, pp. 29-53.

CHAPTER VII

Experiment No. 4

SET AND COMPLEX BEHAVIOR

The importance of set in determining the nature and efficiency of response can be shown not only in the case of the simple reaction, but in practically every type of psychological activity.

Thus the direction of *attention* at any moment is due primarily to a combination of relatively permanent natural sets (instinctive interests) with a temporary set aroused by the particular problem or situation which is faced. It is a commonplace that we notice what we are looking for or what is of interest to us, and often fail entirely to observe the eqully obvious things which do not concern us at the time. So also in *memory* it is clear that the response most frequently, most recently, and most vividly associated with the present stimulus, does not always recur, as the simple law of habit would predict it should, but that recall or recolection often consists of the weakly associated responses which are favored by the set of the moment. The falsification of testimony by a "leading question" which predisposes the witness to give the desired answer without his realizing either the influence of the question or the error of his report, is a well recongized illustration of the influence of set.

In more evidently *"motor"* activities, striking examples are equally easy to find, both in the case of man

and of the animals. After the first few trials in *learning* to run a maze a rat shows that he is set for the goal (feed box) through the fact that he enters considirably more often into "blind alleys" leading towards, than into alleys leading away from the goal. Human learning is most efficient only when there is an intention (set) to learn and to remember.

The "delayed reaction" experiment, which shows that certain animals can choose from among several doors the one over which a light appeared some tine before, provided only that during the delay their mucsular set is maintained, is an example which suggests the importance of muscular sets in the *imagery* of human beings.

Many *illusions* are at least in part due to the observer's set or predisposition. Thus the *size-weight* illusion consists in the fact that of two physically equal weights the smaller one feels heavier when lifted. The subject (whose experience teaches him that larger things are generally heavier) unconsciously *sets* his muscles more strongly for picking up the larger weihgt, lifts it with accordingly greater ease, and therefore receives the impression that it is relatively light.

Together with the "imaginal" or "implicit" character of the processes concerned, the distinguishing characteristic of *thought* is the fact that its course is directed by the subject's set (purpose, intention, determination) to solve some problem. This distinction between unguided and directed "mental discourse" was pointed out by Thomas Hobbes as early as the middle of the 17th century.

Illustrations might be further multiplied. It will suffice here to say that particularly in all forms of "voluntary" or "controlled" mental activity and voluntary

action, the response or sequence of responses is neither simply "random" (or "reflex") nor simply habitual, but determined or directed by the tendencies to reaction, which are established by the instructions, commands, questions, suggestions, or other conditions of the problem or situation. The psychological concept of "set," that is to say, takes the place of the common-sense notions of *"will"* and "intention."

The present experiment is an attempt to measure the influence of definite set, as compared with less definite set, in the problems of discovering words hidden among pied type and in thinking out words from anagrams and skeleton words. The conditions are difficult to control but with careful attention to instructions and technique, fairly consistent and reliable results can still be obtained.

Apparatus.—Paragraphs of pied type containing hidden words. Anagrams for transposition into five-letter words. Lists of five-letter skeleton words. Key lists. (See *Notes for Instructors* for details of the above.) Stop watches or timers for measuring the approximate time of verbal reactions.

Method.—Be careful not to let S read any of the lists ahead of time. Do not talk over results with him, or let S see your records until the entire experiment is completed.

Part 1. *Hidden Words.*—Get from the instructor 8 short paragraphs of pied type. Two paragraphs contain hidden words of unspecified length (I); the other six contain only words of a specified length (2 or 3 or 4 letters). The 8 paragraphs are arranged (to compensate for practice) in the order, I, 4, 3, 2, 2, 3, 4, I.

Instruct S: *"I shall show you paragraphs of pied type, in each of which short words appear in various*

places. Before we start each paragraph, I shall tell you whether the words you are to find and underline are all 2-letter, 3-letter, 4-letter words or are of unspecified lengths. When I give you the signal take the paragraph line by line and try as rapidly as you can to underline all of the actual words. If you overlook very many, I shall send you back and keep you hunting until you discover 10 such words in each paragraph."

Set the hands of your timer just before zero. Tell S: *"In the first paragraph the words are of indefinite length."* Start the timer, and as the second hand crosses zero tell S to begin. Count the number of words he underlines, and stop the timer when he has marked his *tenth* word. (There are 12 words in each paragraph. If S misses more than 2, he must go back.) Record S's total time for marking the 10 words with I set.

Pause and reset your timer. Instruct S: *"In the second paragraph look for and mark 4-letter words."* Stop the timer when he marks his tenth and record the total time elapsed.

Take the remaining paragraphs in order and proceed similarly, giving S the proper definite or indefinite set before he begins each paragraph.

Part 2. (Part 2 and Part 3 of the experiment may be shortened by using stimulus series of 80 instead of 120. The alternate directions in parenthesis refer to the shortened experiment.)

Anagrams.—From the instructor obtain a key list and 6 (4) lists of 20 anagrams each. If they are printed on cards, see that the cards are arranged in order in each set.

Make out 6 (4) tables, each for recording 20 times and 20 responses. Label trials 1-20, I (indefinite); trials

21-40, E (eating); (trials 41-60, H, house; 61-80, H);* 81-100, E; 101-120, I.

Write SAREE on a bit of paper and instruct S: "*These letters, if properly arranged, would make a word. Try to find that word mentally (without use of paper and pencil).*" If he fails to find the word in half a minute show him that ERASE is possible.

Then say: "*I shall show you one at a time a series of anagrams of this sort. As soon as you find a word which uses all of the letters and no more, call it out. Work as rapidly as you can. Try to make real words and avoid slang and nicknames. Never give up any word until I tell you to do so.*"

Give S the ready signal, and as the second hand of the timer crosses zero (or some other easily remembered mark) show S anagram No. 1.

Record responses and the time that elapses before he discovers a word which satisfies the conditions noted. The most common words are on the key list, but any word will do which uses these 5 letters and is not slang or a proper name. Give S credit if you are in doubt. If S *does not succeed in 30 seconds, take the card away and record* "—30."

Similarly proceed with anagrams 2-20, recording the *response* and the *time* taken for each one separately.

Make ready for the next series Nos. 21-40. Tell S: "*The anagrams in this series were made from words which have to do with eating. Set yourself accordingly to make such words. However, if any real word that will do at all occurs to you, give it, and you will receive credit, even if it does not refer to eating.*" Time as before, never allowing over ½ minute per anagram, and

* Omit trials 41-80 in shortened experiment.

record the 20 responses and times of this series in your second table.

(Omit this paragraph in shortened experiment.) Make ready series Nos. 41-60 and tell S: "*The anagrams in this series were made from words that represent things about a house. Set yourself accordingly, but give any real words that occur to you, even if they do not refer to things about a house.*" Then proceed to time and to record the responses as before.

Give appropriate instructions to S before each of the three (two) remaining series, and complete the 120 (80) anagrams.

Part 3. *Skeleton Words.*—Return the anagram lists, properly arranged. Get from the instructor a new key list and 6 (4) lists, each containing 20 skeleton words.

Make out 6 (4) tables, each for recording 20 responses and their corresponding times. Label trials 1-20, I (indefinite); 21-40, A (adjectives); (41-60, I; 61-80, I);* 81-100, A; 101-120, I.

Instruct S: "*These are skeleton words in which two letters of the original five are omitted. When I show you a skeleton word think out and tell me a real word (not slang or a proper name) which could be made by filling in the two blanks with the proper letters.*"

Give trials 1-20, and record S's responses and their corresponding times in your first table. If S does not find a word in 20 seconds, take the card away and record "—20."

Pause and make ready series 21-40. Instruct S: "*This series was made by omitting letters from 20 adjectives. Set yourself accordingly, but if real words that are not adjectives occur to you, give them and you will receive*

* Omit trials 41-80 in shortened experiment.

credit." Record responses and times in your second table, never allowing more than 20 sec. per word.

With appropriate instructions before each, complete the four (two) remaining series of skeleton words.

Return your sets of skeleton words, properly arranged.

Results and Discussion.

Part 1.

(1) Compute the total time taken by your S to find the 20 hidden words with *indefinite* set; to find the 20 hidden two-letter words with *definite* set; the 20 three-letter words; the 20 four-letter words.

Compare your subject's results with the following "typical" results. Point out similarities and try to explain differences:

For 20 two-letter words (def. set), 150 sec.
For 20 three-letter words (def. set), 150 sec.
For 20 four-letter words (def. set), 200 sec.

Average for group of 20 words with definite set, 166 sec.
For 20 words with indefinite set, 266 sec.
A typical subject takes 1.6 ($^{266}/_{166}$) times as long to find 20 hidden words with indefinite set as with definite set.

(2) Did you notice any misspelled words as you read the introductory paragraphs to this experiment? How many? Reread the first five paragraphs again, looking definitely for misspellings. How many do you now find? Ask the instructor to tell you how many are actually misspelled. Did you catch them all even when looking for them? State fully and

clearly in psychological terms (attention, habit, set, etc.) the reason for this "proof-readers' illusion" and for your greater success on second reading. The overlooking of accidental errors in composition (setting up type) by the proof-reader is considered to be largely due to the proof-readers' illusion (the tendency, on seeing a few letters of a word, to fill out the rest from memory and perceive the word as a whole).

Part 2.

In Part 2 and 3, the numbers in parenthesis refer to number of responses used in the abbreviated experiment where the total number of anagrams and number of skeleton words used was 80 instead of 120.

(3) Make quantitative comparisons to show clearly the difference in time taken to transpose anagrams under *definite* and *indefinite* set.

Typical results show that the median time with indefinite set is about 2.5 times as long as for definite set. Compare your findings with the "typical" results and state what the comparisons tend to show. What factor *besides set* may possibly be responsible for the discovered difference?

Part 3.

(4) Go through the two tables (40 responses) for the skeleton words completed under *indefinite* set. After each response that is clearly an adjective mark an "A." Obtain the median time for adjective responses with indefinite set; for non-adjective responses with indefinite set. Compare medians.

Check all adjective responses given under the specific set for adjectives and obtain the median time for these responses. Compare with medians obtained for adjective and non-adjective responses.

(5) Obtain a sheet from the instructor giving order of missing letters in skeleton words. What is the median time for the thirty --xxx responses, i.e., for all responses in which the first two letters were omitted in the skeleton word? Obtain median time for xxx-- responses, for -xxx-, and for x-x-x responses. Judging from these results, which type of omission makes the greatest difficulty? the least difficulty?

Experiments on perception in reading show that letters in certain parts of a word are more important than others for accurate and quick perception of the word. Letters at the beginning are most important, those at the end next in importance, and those in the middle least important. Relate your results (obtained in the first part of this question) to these statements concerning word perception.

(6) In the light of the comparisons made in the 3 parts of this experiment state clearly and concisely the influence of definite versus indefinite set in performing a task.

(7) What bearing has "set" on methods of teaching and study? (Get hints from Dashiell, pp. 370-374 or Pyle, pp. 68-70 for help in answering this.)

(8) Cite controls employed in this experiment and state why they were used. Criticize their adequacy and suggest changes in technique that might improve certainty of results in such an experiment.

Summary and Conclusions.

References

J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 370-374; A. M. Jordan: *Educational Psychology,* 1928, pp. 65-67, 97-102, 132-133, 137-138; W. H. Pyle: *The Psychology of Learning,* 1928 (Rev. Ed.), pp. 62-70; R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 236-240.

CHAPTER VIII

Experiment No. 5

CUTANEOUS SENSATIONS

We commonly speak of touch as one of the "five senses" and suppose the skin to be its origin. In this common broad connotation, touch includes the apprehension of weight, of movement, of roughness, sharpness, wetness, and the like, as well as the simpler sensations of *cold, warmth, contact,* and *pain.*

In reality, only the latter four are simple cutaneous qualities. The others mentioned are complex matters of perception. They usually involve not only cutaneous sense organs, but also organs which are in the underlying tissues—membranes, muscles, fascia, tendons, and joints. Stimulation of organs in the skin alone is hence impossible unless far greater care is taken than is taken in everyday life. Even the four cutaneous sensations mentioned cannot be evoked uniformly over the whole surface of the body, but appear only in certain very small areas, or "spots" arranged irregularly with far larger areas of insensitivity between. Some tissues, such as the cornea and conjunctiva of the eye, the inside of the cheek, etc., are insensitive or nearly insensitive with respect to one or more cutaneous qualities. We fail to notice these facts in everyday life, (1) because we pay so much attention to the *objects* producing the stimulation, and have so little interest in the *sensations* them-

selves; (2) because ordinary stimulation of the skin involves larger areas; (3) because in such cases the whole area involved appears to be "filled out" through summation of the stimuli; or (4) because even moderate stimulation is likely to excite the organs in the deeper tissues and thus to "fill out" the whole area by sensations of similar quality from the deeper tissues.

The cutaneous sense organs are supposed to lie beneath the spots. There is incomplete evidence which indicates that cold is mediated by the bulbs of Krause and of Golgi-Mazzoni, and by the corpuscles of Dogiel; warmth by the Ruffini cylinders; contact by the free nerve endings in the hair bulbs and by the Meissner corpuscles; pain by the free nerve endings in the epidermis.

The purpose of this experiment is to gain subjective familiarity with the cutaneous sensations of cold, warmth, contact and pain; and to map out a skin area showing the location and frequency of spots at which these sensations arise when the skin is stimulated.

Apparatus.—Temperature cylinders, set in corks, pan (for ice or hot water), cloth or paper towel (for drying cylinders), rubber stamp and pad, horse-hair fastened in cork. (See Appendix for detailed description.)

Method.—Have S place his left hand on the table, dorsal side upward. The hand should be lightly closed so that the skin will be slightly strained and smooth. If necessary, clip short or shave off the hairs on this area. Stamp a square on the back of the hand midway between the wrist and finger bases.

Stamp similar squares on small pieces of paper. On these latter, the records are kept. They should be labeled with the name of S and the sides of the squares should be labeled R, L, F and W (for right, left,

fingers, and wrist). The direction of movement of the cylinder point is to be indicated by an arrow.

Part 1. *Cold.*—Chill the cylinders in ice water. Dry a cylinder by wiping with a cloth or paper towel. Have S close his eyes. Move the pointed end of the cylinder

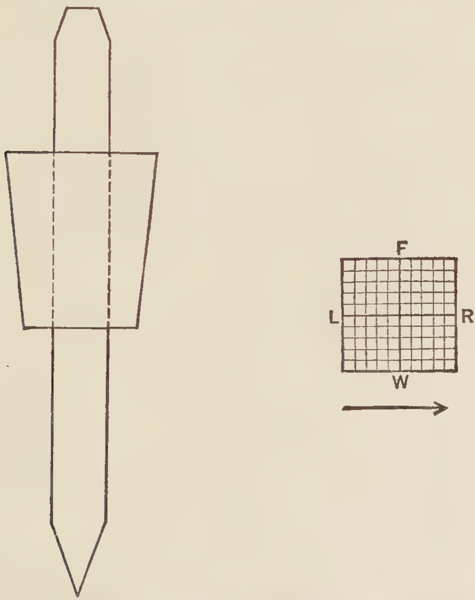


Fig. 10. Cylinder for discovering and map for recording temperature-spots.

steadily and not too slowly over the skin at one side of the stamped area, being careful barely to brush along the skin surface. Pressure organs are very numerous and if too great pressure by the point of the cylinder is used, confusing pressure sensations from the underlying tissues will arise. Be careful also not to touch S

with your hand or sleeve, for such additional contact is likely to distract his attention. Change the cylinders frequently so as to keep a freshly cooled one in use.

Tell S: *"Try to pay attention strictly to the feeling in the skin as I move the point along. Learn to distinguish cold from mere contact. When you feel a cold sensation say 'There,' and if it seems especially clear or intense say 'There—very clear.'"*

Note the location of a few spots of which S feels certain, and some time later move over them again and see if S will confirm his previous locations. It is too much to expect that he will always do so, because the sense organs "adapt" or fatigue so rapidly, but continue preliminary practice until S, by a certain degree of consistency in relocation, shows that he really can distinguish weak cold from pressure.

Now go over the mapped area of the skin. Start at one side and trace across twice between each pair of transverse lines—once near one side and once near the other side of the intervening space. Whenever S says "There" stop and record a dot on the corresponding place on your record map. Record the spots reported as "intense" by tiny crosses.

After a few minutes' rest, repeat the experiment, this time moving across the map in the direction perpendicular to the lines formerly traced. Keep results on a second record map.

Use the blunt end of a temperature cylinder and give S practice in judging the comparative intensity of cold so produced on different areas at one side of the map. He is to report "intense" (I), "moderate" (M), "just perceptibly cold" (J), or "no cold" (O). Then go over in this way 25 areas each composed of 4 tiny squares

on the map, and enter his reports abbreviated as above on the corresponding squares of a record map labeled "cold areas."

Get and note reports from S which will enable you to answer the questions asked later under Results.

Part 2. *Warmth.*—Use cylinders warmed in water which has been heated to above 40° C. Give S a good bit of preliminary practice in locating warm spots. Observation and accurate location are not easy. Move along considerably more rapidly than in locating cold, and change the cylinders more often.

Secure maps in two directions of movement and for areal stimulation on the stamped area, as you did for cold. Get reports which will enable you to answer the questions asked below.

Part 3. *Contact and Pain.*—Try a horse hair projecting about 1¼ inches from a cork. Have S close his eyes. Give "ready," "now" signals and see if sometimes when this hair is pressed down upon his skin till it just begins to bend, S feels the contact and sometimes not. If the hair gives so strong a pressure that it is always felt, lengthen it by pulling it from the slit in cork.

Choose 4 tiny squares at one corner of the stamped area. Bring the hair down four times inside each square, once near each of its corners. Have S report whether "contact" or "nothing" is felt. Record the location of the contact spots found. After a rest repeat the experiments and secure a second record for contact.

Use the sharpened end of the hair. Let it project about ½ inch and see when it is brought down till it just bends, if S will sometimes report cutaneous pain and sometimes "nothing."

Results and Discussion.

(1) Does the cold sensation appear to S larger or smaller than the area of the cylinder point?

(2) Does the cold seem to reside entirely on the surface of the skin? If not, how deeply down into the tissues does it appear to run?

(3) Which of the following adjectives seem to S to be applicable to cold sensations: lancing, blunt, diffuse, compact, dull, bright, lively, sluggish? What other descriptive terms can he think of which seem applicable?

(4) Make a composite map by comparing the maps obtained from tracing in the two directions and by marking down every point which is duplicated within 2 mm. (the width of one tiny square).

(5) How many cold spots were marked on the first map? How many on the composite map? What percentage of cold spots are thus "confirmed"? Is there evidence that intense spots are easier to relocate?

(6) State reasons for failure in relocation in addition to adaptation of the sense organs.

(7) Make a composite map for warm spots.

(8)-(10) Answer questions (2), (3) and (5) as applied to warmth.

(11) Which spots appear more numerous—cold or warm?

(12) Which quality appears more definite and more easily attended to—cold or warmth? Which spots seem larger? Which seem deeper?

(13) Answer question (5) as applied to contact. Be sure to add descriptive adjectives of your own.

Compare contact with tickle; with ordinary pressure given by a pencil.

(14) Answer question (5) as applied to cutaneous pain. Be sure to add descriptive adjectives of your own. Compare this cutaneous pain with itch; with ache; with pains inside the body.

(15) How does the sensation from areal stimulation compare in intensity with that from stimulation of an isolated spot?

(16) Compare the composite maps with the maps from areal stimulation. Can you discover any relationship between number and "intensive tuning" of spots, and intensity reported for the areas including them?

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 316-322; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 84-89; C. S. Myers: *Text Book of Experimental Psychology,* Pt. I, 1911, pp. 10-18; E. B. Titchener: *Text Book of Psychology,* 1911, pp. 143-159; *Experimental Psychology,* Vol. I, Pt. 2, 1901, pp. 81-98; H. C. Warren: *Human Psychology,* 1920, pp. 201-204; B. B. Breese: *Psychology,* 1921, pp. 118-129; W. B. Pillsbury: *Fundamentals of Psychology,* 1922 (Rev. Ed.), pp. 157-170.

CHAPTER IX

Experiment No. 6

COLOR MIXTURE

What we see depends upon the rate of the ether vibrations acting upon the retina and upon their proportions in *color mixtures.* The physical stimuli are, however, but a part of the story. What we see also depends fundamentally upon how long we look (*adaptation*), upon what we have recently looked at (*after-image*), upon what other colors are spatially nearby (*contrast*), upon whether we look directly or "out of the corner of our eye" (*indirect vision*), upon whether our retina is developed completely or not (*color blindness*), upon the intensity of illumination (*daylight, twilight and night vision*) and upon numerous other factors besides. That in practical life we fail to note the importance of all these latter factors is a fact which demonstrates again the difference between scientific observation and ordinary unsystematic notice.

This experiment deals with the first factor and demonstrates the results of mixing visual stimuli. The normal or "adequate" stimuli for vision are ether vibrations at rates between the limits of about 375 trillion and about 750 trillion per second. Under ordinary conditions isolated waves of a single rate only never affect the eye. Sunlight, for example, is composed of waves of many different rates—waves, which if taken singly, would give us literally "every color of the rainbow." Ordinary

"red" glass transmits, and ordinary "red" objects and red pigments reflect not only waves corresponding to various reds, but also waves of some intensity corresponding to the oranges, the yellows, the violets, the blues, etc. The apparent color of all objects (except incandescent objects) depends upon their tendency to absorb some, and to reflect or transmit to the eye others of the waves which make up sunlight.

A strictly controlled experiment in color mixture therefore requires the use of a spectrophotometer or other device for controlling the rate and the intensity of ether-waves. Such instruments are too expensive and not adapted for class use. The general laws of color mixture, however, may be demonstrated roughly by the use of the color wheel and paper discs of impure colors. When in rotation, this wheel presents in very rapid succession now one, now another color. Since the chemical activity of the retina is slow in comparison with the rapidity of alternation of the colors, the total effect is the same as though the colors concerned were simultaneously thrown upon a given area of the retina. The total effect will, of course, depend not only upon the color in the discs but also upon the proportion of each disc which is visible.

The purpose of this experiment is to study the results of mixing visual stimuli (colors). Colors will be mixed to demonstrate the three laws of color mixture: (1) the mixture of complementaries; (2) the mixture of noncomplementaries; and (3) the law of stability (constant mixing value). An unknown color will be analyzed by means of color matching.

Apparatus.—Materials for demonstration of hue, tint, chroma, and contrast. Test material for color blindness.

Color wheel and color discs. Protractor. Milton-Bradley color tops and discs.

Method. **Preliminary.**—The laws of color mixture are stated and the facts of complementarism made more definite by reference to the color triangle, color circle, or other schematic figures.

By use of the spectrum and other charts, the color wheel, etc., the instructor familiarizes the class with the

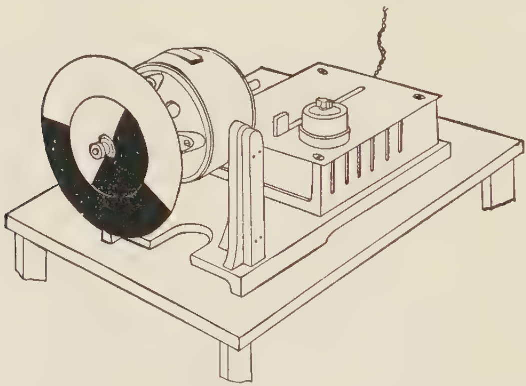


Fig. 11. Motor color-mixer.

distinctions of hue (color-tone), tint (intensity, brightness), and chroma (saturation). Practice in distinguishing tint and chroma and in distinguishing among poorly saturated hues is a valuable preliminary exercise. Students who suspect themselves of color weakness of any sort or degree should be given such a test for color blindness as the Holmgren wools, the Nagel cards or Stilling's pseudoisochromatic tables. It is well also to explain the principle of the color wheel by reference to the positive after-image and to demonstrate the latter. The nature and effects of simultaneous contrast may also be briefly

shown by use of the color wheel, colored papers, colored shadows, etc.

In large classes the following exercises may also be performed as group experiments, the changes being made by the instructor so as to bring about a match, first for some centrally located subject (CS), and then if time permits for individuals in other parts of the class who differ from CS in their judgments.

A table is made in which to record: measurements of the outside disc (3 columns); measurements of the inside disc (3 columns); judgments of CS (3 columns—hue, tint and chroma) and judgments of the individual recorder (hue, tint and chroma). In order to avoid mistakes in recording, an assistant may record measurements and judgments of CS in a large table on the blackboard.

Law I. *The Mixture of Complementaries.*—Begin with 180° of Yellow and 180° of Blue as outside disc and by steps make an approximation to Gray for CS. Then add a smaller disc of 180° Black and 180° White and secure a match for CS, adding a third color to the outside if necessary. All CS's judgments are recorded by all subjects. Uniformity and certain interpretation of the record is secured if all judgments are required in terms of the outside disc as: "too Y," "too B," etc., to match the inside.

Change CS and match R and BG (plus a third color if required) to Bk and W. If time permits solve problems such as "What is the complement of this Violet?" by the same procedure.

Darken the room. Is the match altered? Turn on the electric light. Is the match altered? If changes in

the appearance of matches occur with these alterations of conditions, describe them. Does the match for CS remain a match if he approaches the mixer within 2 or 3 feet? If he recedes from it greatly?

Law II. *The Mixture of Non-Complementaries.*—By the same procedure as in Law I, match an outside disc of R and Y to an inside disc of O, adding white or black or both to equate for brightness and saturation differences when required.

Use R and B in an outside disc and match to an inside disc containing P or V.

Law III. *The Law of Stability* (*Constant Mixing Value*).—Put Y, B, R, and BG in an outside disc in the same relative proportions (½ of each) as in the final match of Law I. In an inside disc use Bk and W in proportions identical with those of the first match (180° of each). Is any change necessary to secure a match? Is any change necessary if the room is darkened?

Color Analysis by Matching.—The instructor will furnish a single unknown or a set of unknown colors for the outside disc. On the inside disc build up from your six primary colors a combination which matches the outside disc exactly in hue, saturation, and brightness. Record the final equations.

Obtain data for questions (3) and (5) below.

Results and Discussion.

(1) Express the matches obtained in the form of color equations.

(2) List as many factors as you can which need to be controlled in order that a match for one person shall be a match for another.

(3) What phenomenon occurs if the color wheel rotates too slowly to give complete fusion? Why? Swing a stick or pencil back and forth between your eye and a rapidly rotating mixture. What phenomenon occurs? Why?

(4) Describe the color triangle and state how complementary colors can be stated by reference to it. What other figures make possible approximate statements of complementaries?

(5) What is the complement of G? of O? of YG? of V? of orange red? of yellowish orange? Use the Milton-Bradley color-top with Bk and W as center, and the color considered in the outside disc. Discover what color or colors need to be added to the outside to secure a match. Express the matches as color equations in terms of the divisions on the scale on the cardboard backer of the top.

(6) Predict as accurately as you can what the results of mixing the following colors will be. ½ Bk and ½ B? ½ W and ½ B? ¼ W and ¼ Bk and ½ B? ¼ P and ¼ G and ½ Bk? ½ R and ½ G? ½ R and ½ YG? ½ P and ½ BG? ½ Y and ½ P? Use the color top with the above mixtures as outside discs and the color nearest your predicted result as a center disc. What additions or alterations are found necessary in order to secure matches?

(7) What colors may be most easily mixed to produce brown? What is the color of grass in the spring? of the illumination from an ordinary electric light?

(8) How did simultaneous contrast help to make judg-

ments surer and easier and matches more accurate in this experiment? Refer to a concrete case.

(9) How do the results of mixtures in this experiment differ from those of painters in mixing pigments? Give reasons for the difference.

(10) Give a physiological theory to account for the general facts of color mixture.

(11) State in terms of the theory what is supposed to occur in the retina when light Y and dark B affect it equally and simultaneously. When Bk and W affect it. When dark R and BG affect it. When light B and Bk affect it. When a small amount of dark B and a large amount of light Y affect it.

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 342-361; E. B. Titchener: *Text Book of Psychology,* 1911, pp. 59-80; 87-92; J. B. Watson: *Psychology from the Standpoint of a Behaviorist,* 1919, pp. 86-105; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 148-162; C. S. Myers: *Text Book of Experimental Psychology,* Pt. I, 1911, pp. 71-101; B. B. Breese: *Psychology,* 1921, pp. 176-196; H. C. Warren: *Elements of Human Psychology,* 1922, pp. 57-84; W. B. Pillsbury: *Fundamentals of Psychology,* 1922 (Rev. Ed.), pp. 109-135.

CHAPTER X

Experiment No. 7

VISUAL ADAPTATION AND AFTER IMAGES

This experiment deals with the influence of time upon the character of visual sensations. You have noticed that when you come from the light into a dark room all at first seems intensely black. If you remain, the black gradually lightens (becomes dark gray) until perhaps certain objects can be distinguished from the general background. If now you again reënter the light it dazzles you (seems intensely white) but it, too, rapidly becomes normal again (gets darker). These are instances of *general adaptation,* the first to black and the second to white. Such experiences as the gradual "getting used" to wearing colored glasses show that general adaptation to color also occurs. Again when a smaller area instead of the whole retina is stimulated, a *local adaptation* takes place. With continued stimulation all light and all colors become more and more like gray.

What we see at any given moment depends not only upon the present stimulus and upon how long we have been looking, but also upon what we have just previously looked at. Stimulation of the eye, that is to say, leaves after effects which in some cases may persist for a considerable time. The most easily observable after effects are the *positive after-images.* An instance from daily life is the bright patch or form which stays before our eyes after we have looked directly at the setting sun or

at a very bright light. Still more definite positive after-images are obtained by brief but intense stimulation after prolonged dark adaptation. In such cases the positive after-image may also be followed by a "flight of colors."

When local stimulation is prolonged and not too intense and the subject afterwards gazes at a neutral gray ground, a *negative after-image* of a color approximately complementary to that of the original stimulus is seen. The after effect of general adaptation under the same circumstances is a contrary (complementary) "sightedness." Thus everything seems more bluish when first we come out of the theater from a late matinée because during the performance we were steadily subjected to the somewhat yellowish artificial light.

The phenomena of adaptation and after-images are noticed so little in daily life, first because we pay attention to objects and their practical aspects rather than to colors in and for themselves; secondly because our eyes are so constantly in movement that strong adaptation and after-images are rarely produced; and thirdly because the changes due to adaptation take place so slowly as to make comparative observation difficult. Some subjects have difficulty at first even under laboratory conditions, chiefly for the first reason mentioned—they do not yet realize exactly what to look for. Once the phenomenon is observed, however, it is afterwards unmistakable, though further practice in fixation and in verbal report may increase the accuracy of observation and of description.

The purpose of this experiment is to become subjectively familiar with and describe: (1) adaptation to

visual stimuli, (2) positive visual after-images (after sensations), and (3) negative visual after-images (sensations).

Apparatus.—Strong electric light in dark room. Oblongs and squares of black, white, gray and colored paper or cardboard. (See Appendix for details.)

Method.—Dark adaptation, positive after-images, and the flight of colors may be demonstrated by the instructor to groups in the dark room. After dark adaptation for about five minutes a single stimulation of 1 sec. by the electric light will produce a fairly prolonged positive after-image if the light is directly fixated. By a succession of shorter stimulations, positive after-images of directly fixated white objects, and even of nearly all details in the room (the so-called "ghost images") may be produced with rather startling definiteness and "reality." A 3 to 5 sec. stimulation should give a good flight of colors. Students are required to record notes on these demonstrations immediately after leaving the dark room, in order that they may answer fully and accurately the relevant questions asked under "Results."

The following observations of adaptation and the negative after-image are to be performed by all members of the class individually. In describing a color, report its hue (color tone), its tint (brightness), and its chroma (saturation), *all three.* Do not look away from an after-image as soon as you think you "have" it, but watch its full course, and describe it *fully.* Success depends upon absolute obedience to instructions, concentration of attention, steady fixation, and painstaking and careful observation. Close the eyes for rest between experiments as much as possible. Do not allow them to wander aim-

lessly over colors not in actual use. Sit back to the light so that there are no shadows cast at the junction of papers when one is placed over another.

(a) Cover one-half of a large Bk oblong with a gray one. Fixate an ink dot at the junction, and hold your eyes steadily upon it while you count 100 slowly to yourself. The gray changes little or not at all, but slow changes take place in the appearance of the Bk. Notice the latter "out of the corner of your eye," i.e., without removing fixation from the ink dot.

What took place on the gray near the junction? On the Bk near the junction?

After such long steady fixation suddenly remove the gray and hold fixation steadily on the center of the Bk oblong. You now see one-half of the Bk where adaptation has taken place, and one-half of it (that covered by the gray during fixation) which has not been subject to adaptation and therefore looks as the Bk originally did. Compare the adapted with the unadapted half.

(b) Substitute a W for the Bk oblong used in (a), cover one-half with the gray, and repeat the observation.

(c) Substitute in succession for the W or Bk, oblongs of the following colors: R, Y, G, B, and repeat the observations. Do not fail to report hue, tint, and chroma in all descriptions.

(d) Make out a tabular form. Label the first column "Stimulus." Divide the remainder of the table into 3 wide columns labeled "Gray Bkgrd," "Bk Bkgrd," and "W Bkgrd." Subdivide each of these 3 wide columns into 3 narrower columns, entitled "Hue," "Tint," "Chroma."

Place a small W square in the center of a large gray oblong. Fixate the ink dot in the center of the W while

you count 50. Suddenly remove the W square and observe the negative after-image by gazing steadily at the ink dot in the center of the gray background. Record the stimulus and the results of your observation in the first row and in the appropriate columns of the table.

Repeat, using the W square and a Bk background. Then use the W square with a W background. Fill in the appropriate columns of the first row.

Now use a Bk square as stimulus, and gray, Bk, and W in succession as projection backgrounds. Record observations of the after-images in the second row of the table. (Some columns remain vacant until colors are used as stimuli.)

Repeat again, using R as stimulus and projecting the after-image on gray, Bk, and W backgrounds and keeping a record in the third row of the table.

Repeat, using Y as stimulus, then G, then B, then V, and recording the after-image of these colors when projected on the gray, the Bk, and the W backgrounds.

(e) Make a table similar to that used in (d), but with 4 triply subdivided columns for recording the hue, tint and chroma of after-images projected on R, Y, G, and B backgrounds.

Use small squares of W, Bk, R, Y, G, B, and V in succession as stimuli and record in the table the colors of the after-images when projected on each of the 4 colored backgrounds.

(f) Place a small square of W paper in the center of a large Bk oblong. Fixate the center of the W steadily while you count 50. Close your eyes and put your hands over them. Describe the appearance and course of the after-image in the darkened field of vision. What is its apparent size? its apparent distance? Does it move? If

so, explain why. Repeat, using a small square of Bk and a large square of W.

(g) Place a small square of W or of some color on a large Bk oblong. Fixate until you can get a strong negative after-image. Project this after-image first on an oblong quite near your eyes, then on a nearby wall, then on a wall across the room. Compare the apparent size of the after-image in the three cases. Project on a wall not perpendicular to you, and in a corner where two walls or walls and ceiling meet. Is there any alteration from the "normal" shape of the after-image?

Results and Discussion.

(1) Why does intense stimulation favor positive after-images? Why does previous dark-adaptation favor positive after-images? Why do positive after-images tend to "drift" or move slowly across the field in some direction in the dark room? Get a positive after-image and project it on the wall in a light room. Does it tend as strongly to drift?

(2) What was the sequence in the "flight of colors"?

(3) Describe the "ghost images."

(4) What is the *general* law of adaptation? Summarize the results of exercises (a), (b), (c) above.

(5) What is the general law for the color of negative after-images when projected on a gray background? What is the effect of a bright (W) background? of a Bk background? Summarize the results of Exercise (d).

(6) What is the reason for the bands seen along the junctions in Exercises (a)-(g)?

(7) State the general law of the effect of color of projection background upon color of negative after-

images so clearly that one can use the law to predict what the after-image will be in case of any given stimulus and background. Summarize the results of Exercise (e).

(8) Summarize the results of Exercise (f).

(9) Summarize the results of Exercise (g). What is the general law of the effect of distance of projection background upon size of after-image? Give reasons.

(10) Distinguish between so-called "successive" and "simultaneous" contrast.

(11) What may be supposed to take place in the retina during adaptation to B? to YG? during the ensuing negative after-images?

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 345-346, 355-357; E. B. Titchener: *Text Book of Psychology,* 1911, pp. 71-78, 87-92; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 156-158; B. B. Breese: *Psychology,* 1921, pp. 94-95, 181-186; W. B. Pillsbury: *Fundamentals of Psychology,* 1922 (Rev. Ed.), pp. 109-134.

CHAPTER XI

Experiment No. 8

THE VISUAL PERCEPTION OF SOLIDITY AND DISTANCE

The list of factors known to play a part in the perception of tridimensional space is a long one, and only a brief mention of some of the most important factors can here be given. For more extended discussion the references below should be consulted.

(1) A certain indefinite "*voluminousness*" (extent and tridimensionality) seems to be given immediately in the visual qualities themselves. Partial appreciation of this inherent spatiality can be gained by observing the blackness in a totally dark room or the field of vision with the eyes closed. The blackness seems spread out and thick, without at the same time having definite size or boundaries or outlines or depth, and without being located at a definite distance from the observer.

(2) The voluminousness inherent in sensations other than visual ones can be directly or indirectly compared with that of vision. Such voluminousness is clearly noticed when incompatibility with vision is felt. The greater size of a tooth cavity as appreciated by the tongue, early experiences in failing to reach, grasp or handle objects, the long time taken to walk to an apparently nearby mountain, are examples of such incompatibility. *Spatial experiences with other sense departments*, therefore, appear to assist in "visual" apprecia-

tion, as well as to work towards the greater definiteness and stability of all space-values.

(3) The acuity of vision is greatest near the fovea, and very much less on the outlying portions of the retina. The reflex tendency of the eyes to move so as to bring the images of objects attended to upon the area of clearest vision is therefore of considerable importance as a cue. In particular, the patterns of kinæsthetic sensations arising from the twelve muscles which surround and move the two eyeballs in the process of *convergence* differ perceptibly for lesser and for greater distances of objects which are less than 20 meters away, and thus serve as cues to the distance of nearby objects.

(4) The images of objects are brought to a focus by adjustments of the lens made by the ciliary muscle. These movements of *accommodation* appear to be a cue to distances of less than 2 meters.

(5) Since the two eyes are set apart in the head (about 6½ centimeters from center to center) the images formed on their respective retinas, although they overlap and are identical for certain points, nevertheless differ distinctly for near objects and perceptibly for much more distant objects. Most objects not directly fixated therefore give *double images,* which, although not in themselves commonly observed, are especially important factors in the perceptions of solidity and distance. This *disparity* of retinal images is in theory a possible cue for distances up to 2700 meters (about 1.6 miles).

In practice the importance of double images in the perception of distant objects is overshadowed by that of a number of "secondary" or less "immediate" factors. The most important of such secondary factors are: (6) *aerial perspective,* the increased dimness and blue-

ness of objects at greater distances which make us misjudge distances in fogs or in exceptionally dry atmospheres; (7) the apparent *size* of known objects, with the special ease of *linear perspective* illustrated by the apparent narrowing of the parallel lines of a railroad track and the apparent alteration of the right angles of a table when seen from an angle; (8) *interposition,* the covering or shutting from view of a part of one object by another nearer object; (9) the *configuration of shadows* and the *distribution of light and shade,* a cue constantly depended upon by artists to suggest solidity, depth and distance; (10) the apparent *movement* of objects with movements of the head or eyes, illustrated by the apparent directions of movement in the landscape watched through the window of a moving train.

The purpose of this experiment is to endeavor to isolate (so far as is readily possible) certain of the factors involved in the perception of solidity and distance and to demonstrate their nature and importance as cues. This will be done by studying double images, monocular vs. binocular estimation of distance, and stereoscopy.

Apparatus.—Two sticks inserted upright in wooden blocks. Three pins in corks. Background screen. Occlusion screen. Scissors. Ruler. Brewster stereoscope. Wire model of frustrum of a pyramid. Set of Titchener's stereoscopic slides.

Method.

Part 1. *Double Images.*—Choose a high table or shelf whose top can be brought to the level of S's eyes. Arrange the two upright sticks upon it with the near one about 8 inches, and the far one about 2 feet from S's face. Be sure that the sticks are directly in front of him

(in the median plane which cuts through his nose, and not merely lined up for his favored eye). Be sure that S holds this position and does not move his head from side to side.

Have S fixate (focus upon) the near stick and report the appearance of the far one. Then have him fixate

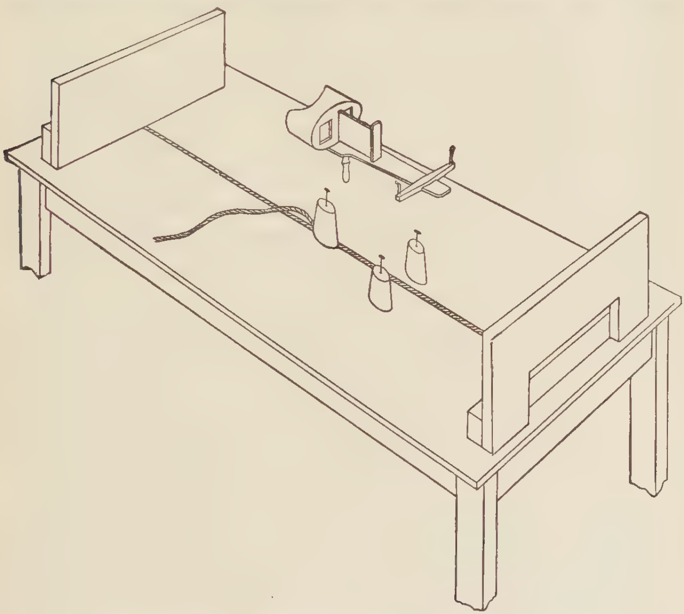


Fig. 12. Stereoscope and arrangement of apparatus for measuring the accuracy of the perception of distance.

the far stick and report the appearance of the near one.

Let S repeatedly try each fixation, and after fixation, close and open now one eye and now the other. Which side (right or left) of the double image disappears when the *near* stick is fixated and the *right* eye closed? When the *left* is closed?

Which side of the double image disappears when the *far* stick is fixated and the *right* eye closed? the *left* eye closed? .

Have S fixate the far stick with both eyes open: Move the near stick still nearer to him, and then further from him. Have him fixate the nearer stick while you move the farther one back and forth. How do the double images change with distance?

Part 2. *Monocular vs. Binocular Estimation of Distance.* Seat S about 5 feet from a high table and so that he can bring his eyes down to the level of the table top. On the table edge perpendicular to S's line of vision set up the occlusion screen. Place two corks with pins about 3 inches apart in a line parallel to the screen and about a foot behind it. Several feet back of the pins set up a dark background. Make sure that S can see the pins through the slit in the occlusion screen, but that he *cannot see the corks, the table top, nor the heads of the pins.* Stretch a paper millimeter scale or lay a meter stick on the table top between the corks, and perpendicular to their line. On this scale a third pin in a cork with a wire handle is to be moved. Make sure that S cannot see the head of this pin, nor the moving cork from which it projects.

Instruct him: "*Watch without moving your head, and tell me to stop when you judge that the middle pin is in line with the other two.*"

Begin with the middle pin about 10 cm. behind the other two, and move it slowly and steadily along the scale till S tells you to stop. Record the error (distance of the middle pin from the line of the stationary pins). Begin 10 cm. in front of the pins and move along the scale in the opposite direction until S stops you. Make

20 such trials (10 in each direction) and record the errors.

Repeat the above experiment allowing S to use *one eye* only, and making sure that he does not move his head while watching. Record the errors in a second table labeled "Monocular."

Repeat the experiment with monocular vision again, but tell S: "*As I move the middle pin, move your head somewhat from side to side, and allow this to assist you in forming your judgment.*"

Part 3. *Stereoscopy.*—(a) Hold the frustrum of a pyramid directly in front of and about a foot from S's nose, on a level with his eyes, and with a small end of the frustrum pointing directly towards him. Have S close his left eye and draw in outline what he sees with the other. Accurate proportions result from first noting carefully the position of the small square relative to the large square base as seen from this point of view. Draw a similar view as seen by the left eye, the right being closed.

Copy these two drawings carefully on a piece of card-board or paper the size of a stereoscopic slide, using the ruler to give straight outlines. Have the center of the two drawings between $2\frac{1}{2}$ and 3 inches apart, and the drawing as seen by the right eye on the right side of the slide.

View this card in the stereoscope and describe what is seen.

Hold the card about a foot in front of the eyes of S and have him practice fixation on a point *nearer* him than the card. By proper fixation two of the double pictures may be made to fuse. Describe the perception of the fused images.

Have S try persistently to fuse the images by focusing *beyond* the card and similarly describe the results. Explain.

(b) Hold the frustrum directly before S, with the base towards and the small end away from him. Have him sketch the view given the right eye and the view given the left eye separately, make a second slide, and describe the appearance when viewed in the stereoscope. Compare this slide with the one formerly made.

(c) Examine in order a group of 10 slides of Titchener's series. In each case, before looking through the instrument, have S try to predict what will later be seen through the instrument. Do not let S merely "guess," but get him to justify and explain his predictions. Keep parallel records of predictions and descriptions of figures seen.

(d) Have S observe in order the group of slides, the two sides in each of which are so different in outline or color that steady fusion is not to be expected. Have him describe the resulting phenomenon of retinal rivalry.

Results and Discussion.

(1) Summarize the results of the observation of double images with the two sticks. Show fully and clearly how double images may serve as cues to distance. Illustrate by diagrams or sketches.

(2) What factor or factors in the perception of distance are eliminated or minimized by using pins instead of larger objects? By using an occlusion screen? By using one eye only? By not allowing movements of the head? By not allowing S to see the heads of the pins?

(3) What was the average error of judgment of dis-

tance with binocular vision? With monocular vision with head stationary? With monocular vision with the head moved? What factors do you consider were the chief ones involved in each of these three cases? What factors do you consider chiefly responsible for the differences in the average errors discovered in the three cases?

(4) Under ordinary conditions judgments of distance are almost as good with the single eye as with both. List all the factors which may be operative in monocular vision. What factor or factors are absent? How may the absent factors be compensated for?

(5) Why do connoisseurs aid themselves in getting the full effect of a painting by viewing it with one eye through a tube made by the hollow hand?

(6) Describe the principles of the Brewster stereoscope.

(7) Suppose that you view an object directly with the right eye, while to the left eye is brought an image of that object taken from a point at the right of your head (as in Stratton's pseudoscope). What appearance should result from the fusion of the two images if the factor of double images is stronger than that of all others?

(8) In what percentage of cases did S predict correctly in part 3, exercise (c)? What rule can be given for prediction?

(9) What are the chief descriptive characteristics of retinal rivalry? Under what conditions does retinal rivalry take place? What happens when black rivals white? When one color rivals another?

(10) What errors in perception would a man born blind, who recovers vision by operation, be likely

to make? Some students are amazed to learn that the images of objects are upside down upon the retina. Would learning to perceive up and down, right and left, far and near, be any easier for a child if the image were right side up? Justify your answer by reference to experiment, and explain how such perceptions develop.

(11) Cite controls used in this experiment, criticize their adequacy, and suggest changes to improve the experiment.

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 357-361, 397-402; E. B. Titchener: *Text Book of Psychology,* 1911, pp. 303-332; G. M. Stratton: *Experimental Psychology and Its Bearing upon Culture,* 1914, pp. 142-151; B. B. Breese: *Psychology,* 1921, pp. 212-222; C. S. Myers: *Text Book of Experimental Psychology,* Pt. I, 1911, pp. 262-272; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 409-413; H. C. Warren: *Human Psychology,* 1920, pp. 234-247; W. B. Pillsbury: *Fundamentals of Psychology,* 1922 (Rev. Ed.), pp. 298-325.

CHAPTER XII

Experiment No. 9

THE MÜLLER-LYER ILLUSION *

The Müller-Lyer figure is named for an early investigator, F. C. Müller-Lyer, who published in 1889. It appears in a great variety of forms which differ one from another in minor details (see references below). In its typical and most convenient form for experiment, it consists of two lines, with "arrow heads" at the extremities of one and "arrow feathers" at the extremities of the other, thus: <——>——<. The first line is of definite length, but the other is so arranged that its length can be varied at will.

The importance of the figure for theory is evidenced by literally hundreds of experiments which have been performed to determine the dependence of the illusion upon form and color, width and length of primary and secondary lines, size of angles, definiteness of the oblique lines, etc., in the figure; upon the naïveté, culture, practice, suggestibility and expectation of the observer; and upon his direction and concentration of attention, his eye movements and his general tendency to "read himself into" the figure (empathy).

These experiments have thrown light not only upon the "explanation" of illusions, but have greatly increased our general knowledge concerning the factors

* It is advisable for the instructor to illustrate to the class several other illusions and to discuss causes of illusions as an introduction to this experiment.

which are important in our normal or usual perceptions of space. The illusion is also of interest from the practical point of view, since under many circumstances its amount is so considerable that designers, draughtsmen, architects and artists need to take account of it.

There are a number of reasons for including this experiment in the course.

(1) The figure produces a so-called "optical illusion," that is to say, an immediate impression which disagrees with the results of physical measurement. The experiment, therefore, illustrates strikingly the difference between the "dependent" (psychological) and the "independent" (physical) aspects of our spatial experience, and shows how greatly mental processes depend upon the points of view, attitude and bodily reactions of the observer.

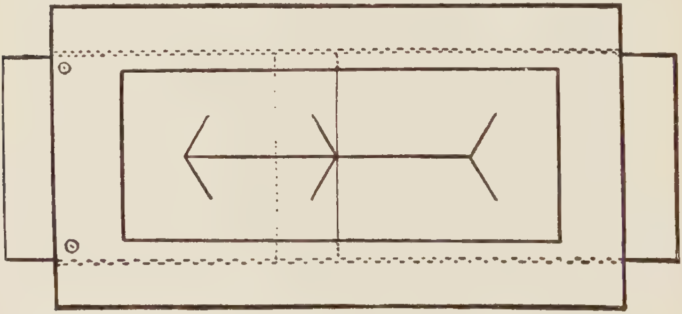


Fig. 13. Illusion board.

(2) The illusion lends itself readily to definite experimental control and the principles of method involved will find extended and general application in many other experiments.

(3) In particular the experiment demonstrates the

variability of the organism due: (a) to "*progressive*" factors (or "errors") such as practice and fatigue, (b) to relatively "*constant*" factors, such as changes in the external stimulus and in its mode of presentation, and (c) to "*accidental*" factors such as involuntary shifts of attention and interest, distractions, inaccuracies of muscular adjustment, and the like. It shows the possibility of *measuring* the influence of such factors by *systematic serial variation* of conditions.

The purpose of this experiment, then, is: (1) to illustrate the quantitative measurement of an optical illusion showing the dependence of the extent of illusion on the physical aspects of the stimulus; (2) to demonstrate how the illusion varies with practice, fatigue, mode of presentation, involuntary shifts of attention and other accidental factors. Controls will be illustrated by means of the ABBA sequence of series.

Apparatus.—Illusion board with two sets of stimulus-cards. See Appendix and *Notes for Instructors* for description and for details of construction. In Set A the obliques lie at a 15° angle to the central line; in Set B, the obliques are at a 60° angle. Measurements are made on a millimeter scale. Thumb-tacks or clips hold the invariable slide in position.

Method.—Note the number or other designation of the particular board you use. Small differences in models may exist, and in case you do not finish experimentation in one laboratory period it is important that you use the same model in completing the experiment.

Prepare 16 separate two-column, five-row tables. Label the first column "Trial No." and the second "Meas." In the first column number consecutively from 1 to 80. The second column is for a record of the measurements in millimeters.

Entitle each table according to the following scheme, which designates the constant condition under which each series of trials is to be taken. A indicates that the cards of Set A are to be used; B designates Set B. R indicates that the variable side of the figure is to be on S's right; L, on S's left. O indicates a preliminary setting such that S will move the variable slide outwards to equate the sides of the figure; I that S will move it inwards to equate:

Trials	1-5	A R O	Trials	41-45	B R O
"	6-10	A L O	"	46-50	B L O
"	11-15	A R I	"	51-55	B R I
"	16-20	A L I	"	56-60	B L I
	2 min. rest			2 min. rest	
Trials	21-25	B L I	Trials	61-65	A L I
"	26-30	B R I	"	66-70	A R I
"	31-35	B L O	"	71-75	A L O
"	36-40	B R O	"	76-80	A R O
	3 min. rest			Supplementary series of 5 trials	

The series order follows the ABBA sequence. Notice also in the above schema that the other conditions, R *versus* L, and O *versus* I, are also arranged on the ABBA principle, and notice that the effects of progressively increasing habituation and fatigue or progressively changing interest will tend, like practice, to be distributed so as not to obscure the effects of other factors. It is now easy to see how by proper combinations and comparisons of the measurements taken under the different conditions, reliable measurements of the effects of each separate factor can be made.

Before each series, arrange the board as prescribed. Make sure that the stable slide is thumb-tacked in such a position that the variable slide can be pushed in or pulled out to a position of apparent equality without interference of S's finger with the end of the board, and with the whole figure fairly well away from the borders of the illusion board.

Seat S with his back to a window, so that no shadow falls where the slides join.

Instruct S: "*You are to hold the board squarely before you at about reading distance, and gradually and steadily draw the adjustable slide in or out until the two horizontal lines appear equal in length. Make your judgments promptly. Do not 'calculate' or make deliberate correction for the known effect of the illusion, but simply move the slide slowly until the two lines 'look' or 'appear' equal in length.*"

After adjustment S returns the board to E, who measures and records in millimeters in the appropriate table the *length of the variable horizontal.* Carefully prevent S from getting any hint as to the nature of any of his records until the eightieth trial has been completed. And in particular until the end, NEITHER E NOR S MUST MEASURE OR KNOW THE LENGTH OF THE INVARIABLE HORIZONTAL. Take the series of 80 trials as a mechanical, impersonal task.

Note and record: (a) whether S is *right or left handed;* (b) whether S has any known *visual defect* or inequality of the two eyes; (c) any *remarks* or other behavior of S which will indicate great or little interest, care, expectancy, fatigue or boredom, etc., on his part at any time. Such notes of S's remarks should be made by E without comment. [S's "natural" or volunteered remarks often throw light upon mathematical results,

and may explain "peculiarities" in his results.] (d) a description of S's *method* of adjusting, especially any definite "method," "tricks," or "mannerisms" that he uses, and any evidence to show that S "calculates" or "makes allowance" for the known effect of the illusion; and (e) the *approximate* time needed by S to complete 80 trials—this to show whether S is performing the experiment more rapidly or more slowly than the average S.

At the end let S know the amount of illusion he has been subject to. Then repeat *the last* sub-series (ARO).

Results and Discussion.

(1) Summarize the notes on points (a), (b), (c), (d), (e) above.

(2) In each sub-series, compute the average amount of illusion, i.e., subtract each of the individual measurements from the length of the invariable horizontal (recording these amounts as negative whenever they so appear), obtain the *algebraic sum* of these errors in each sub-group and divide by 5 for the average.

What is the average amount of illusion for your S with a 15° angle and under our particular conditions of experimentation (take average of 8 averages involving Set A). Compute average amount of illusion with 60° angle (Set B). Compare the two averages.

The illusion with Set A is found greater by about 97 per cent of students. Try to give a verbal explanation of why it should be so. See Myers' reference below.

(3) How does practice appear to effect the amount of illusion? Compare average of first 40 trials with

average of last 40, and average of first 20 with that of last 20.

Practice decreases the amount of illusion for about 80 per cent of students. What explanation for the apparent effect of practice upon the amount of illusion can you offer? See the first Myers' reference below for help.

(4) In what direction and how greatly do "moving in" and "pulling out" (expectation) effect the amount of illusion? To answer, compare O with I series. About 90 per cent of students find the illusion greater with condition I.

(5) How and how greatly does position right or left (space error) appear to effect the amount of illusion? (Compare R and L series). With large groups the differences are small. About 50 per cent find the average amount of error greater on the right and 50 per cent greater on the left.

(6) What effect has knowledge of the actual length of the invariable and of the amount of illusion on later results? Answer by comparing supplementary series with last sub-series.

(7) Suppose that for an unpracticed S, E wishes to compare the reaction time to light with reaction to noise. Reaction time, let us say, depends not only upon the stimulus, light or noise (L or N) but also upon practice, and upon 2 sorts of instructions, "muscular" and "sensorial" (M and S), 50 trials to constitute a sub-series. Make out a schema to show the order of sub-series for the method of *systematic serial variation* (ABBA principle) for this experiment.

(8) State in a few words by what means the method of *systematic serial variation* takes account of "acci-

dental errors'' (factors). Of ''progressive errors.'' Of ''constant errors.''

(9) What are the general explanations of illusions? Illustrate. See Woodworth or Gates reference below.

What do you consider the best explanation or explanations of the Müller-Lyer illusion? Explain why. What other factors besides those we have considered may be important in this illusion? See Myers' reference below.

Summary and Conclusions.

References

For variant forms, see E. C. Sanford: *A Course in Experimental Psychology,* 1903, pp. 212-260, esp. p. 231; E. B. Titchener: *Experimental Psychology,* Vol. I, Pt. 1, 1901, pp. 151-170.

For general explanations of illusions, see R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 412-420.

For experimental results and their bearings upon ''special explanations'' for the Müller-Lyer illusion, see C. S. Myers: *Introduction to Experimental Psychology,* 1912, pp. 45-61; *Text Book of Experimental Psychology,* 1911, pp. 285-292; E. B. Titchener: *Experimental Psychology,* Vol. I, Pt. 2, 1901, pp. 321-328; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 389-405; W. B. Pillsbury: *Fundamentals of Psychology,* 1922 (Rev. Ed.), pp. 329-337.

CHAPTER XIII

Experiment No. 10

SPECIAL PHENOMENA OF VISION

I. *Visual Tests and Color Vision*

While correct vision is important in all walks of life, it is especially so in the educational field and other callings where much reading and other fine visual work are required. In emmetropic or normal vision the image of the object looked at is brought to a sharp and clear focus upon the retina of the eye. This is accomplished by a change in shape of the lens located in the fore part of the eye. The change is called accommodation. For focusing on near objects the lens becomes thicker and more curved while for far objects it becomes thinner and flatter. Defective vision is not infrequent. In hyperopia or farsightedness far objects are seen clearly but usually the lens cannot be altered enough to bring near objects to a sharp focus. The near objects are focused somewhat behind the retina due to a too short diameter of the eye from front to back which causes a blurred image and unclear vision. In myopia or nearsightedness near objects can be seen clearly. For far objects, however, the rays of light are converged before they reach the retina due to a too long diameter of the eye from front to back. This results in a blurred image. Both of these defects are readily corrected by artificial lenses (glasses). Astigmatism is caused by an irregularity of curvature in either the cornea or the lens.

This irregularity causes a blurring of the optical image. Artificial correction is possible. These common forms of visual defects can often be detected by simple tests.

Another kind of visual defect is *color blindness.* Two principal forms of color blindness are: (1) Total color blindness, a form which is very uncommon. An individual who has this defect is able to see only black, white, and shades of gray. (2) A more common form is red-green color blindness. Persons with this deficiency can see blue and yellow along with black, white and gray. About 4 per cent of all males and ½ of one per cent of females are blind to red and green colors. Defects in red-green color vision vary from mild deficiency (color-weak) to complete blindness for red and green. Color blindness is of considerable practical importance. Many of the highway traffic signals, navigation and railroad signals are red or green lights. Confusions arising from color blindness might lead to misinterpreting signals in such situations and produce accidents. Also it is very difficult or impossible for colorblind people to match colors accurately as in selecting dress goods, or to put colors into harmonious combinations. Every individual should have a reliable test for color blindness sometime early in his school life. Color blindness is usually congenital and cannot be corrected.

Everyone is colorblind in certain regions of the retina. The normal retina of the eye is not uniformly sensitive to color over its whole surface, but consists of three distinct zones or regions of sensitivity. Over the innermost zone which is the most efficient of the three are seen all colors (red, green, blue, yellow, and intermediate hues), black, white, and gray; over the middle zone are perceived yellows, blues, black, white, and gray; and over the outer-most zone, which is colorblind, and whose outer

limit corresponds to the edge of the field of view, are seen only white, black, and gray. The limits of these zones are not absolutely fixed but vary somewhat with different intensities of stimulation.

The problem in this experiment is: (1) to demonstrate the use of certain tests of visual acuity and tests of color blindness, and to note the significance of results obtained from such tests; and (2) to determine the limits of the color zones along a single axis of the retina.

Apparatus.—Snellen's Test Type (or Lowell's Type), illiterate E-test type, astigmatic chart, the Ishihara tests for color blindness (optional—Stilling's Tables for testing color blindness), the Holmgren Worsteds for testing color blindness. A perimeter (describe it) with colored stimuli.

Method.—This is a lecture-demonstration experiment.* The instructor acts as experimenter and the class or certain members of the class as subjects. The students are to take notes for use in their reports.

Part 1. *Tests of Visual Acuity and Astigmatism.*—Show these tests to the class, explain how they are employed, and how scored. (If desired the instructor may demonstrate their use by selecting subjects from the class.) For testing visual acuity hang the Snellen chart on the wall about 4 feet from the floor and in good light. Place the student 20 feet from the chart. Care should be exercised that the glare of the sun does not fall upon the chart nor upon the student's eyes. Have the subject hold a card before one eye without touching the eyeball while the other is examined. Instruct the student to read aloud from left to right the symbols in the line marked 20 feet (indicated by the instructor). If

* The instructor will find it necessary to expand both the introduction to the experiment and explanations.

the subject reads correctly 50 per cent or more of the letters in the line that eye is scored R(L) $^{20}/_{20}$, which is normal. The numerator stands for the distance of S from the chart and the denominator for the smallest line that can be read satisfactorily. If the 20-foot line is not read satisfactorily but the 30-foot, 40-foot, or 50-foot one is, then the score becomes $^{20}/_{30}$, $^{20}/_{40}$, or $^{20}/_{50}$ respectively. Any score that yields a fraction equivalent to less than 1 denotes less than normal visual acuity. If, on the other hand, the 10- or 15-foot line is read the subject is probably farsighted. If any abnormality is discovered, an oculist should be consulted.

The E-test is used and scored just as the Snellen chart except that the subject, in reading the symbols, states in each case whether the prongs of the letter point right, left, up, or down.

Hang up the astigmatic chart as in the previous tests. To use this test, the subject covers one eye with a card, looks at the chart and notes whether any of the radii appear darker than others. To the normal eye all radii are equally dark but to the astigmatic eye some appear considerably darker than others. This test is not as efficient as the visual acuity tests.

Part 2. *Color Blindness.*—Lay out on a table or place against a blackboard the Ishihara test so that plates 1 to 13 are visible. Furnish each member of the class with a piece of paper and have the students file past the test plates and write down in order the numbers read on each plate. Inability to read the number on any plate should be indicated by a dash in the series. Then the instructor reads off the responses that a normal subject should make. If any member of the class has more than one response in his series which is different from the in-

structor's reading, tabulate all of that person's responses on the board. If more than one S is abnormal tabulate his responses also. In case no colorblind individual is present record on the board the following as the typical responses of a totally red-green colorblind person: 12, 3, 5, 2, 21, —, —, —, —, —, —, —, -2. Also record the responses of one normal S on the board for comparison.

Show Stilling's Tables to the class and explain their use. Stilling's Tables may be used instead of the Ishihara test since they are very similar.

Now let one normal S and the colorblind S sort the Holmgren Worsteds according to directions with the test. Have the class take notes on this sorting in order to compare performances of the two subjects. They should note especially the number of color confusions and the hesitancy on the part of the colorblind individual. Both the Stilling's Tables and the Ishihara test are valid tests of color blindness but the Holmgren test is not.

Part 3. *Distribution of Color Sensitivity Over the Retina.*—The instructor acts as E and one normal (not colorblind) member of the class as S. Have S sit so that he cannot see the blackboard where E records results for the students to copy for their computations. The apparatus used is the perimeter. This is adjusted with the arm horizontal and at the exact level of S's eye. Have S sit close to the table where the apparatus is located, chin on support, and right eye focused on the fixation mark which is a bit of white paper or a hole at the center of the apparatus. The other eye is held closed with the left hand. The instructor now places a stimulus card carrying a colored dot on the rider, and keeps the card covered with his hand until the rider is adjusted

at the end of the arm. After warning S to hold his fixation, E moves the rider *very slowly* toward the fixation point until S reports that he first sees the stimulus dot but does not recognize the color. Have S close his eyes while E records, from the scale on arm, the angle in degrees at which the dot was first recognized. Place this record under I in column, "yellow" opposite trial 1, if yellow is the color used. E covers stimulus while S adjusts fixation and then E continues to move the rider slowly toward the fixation point until S recognizes and names a color. If the named color is incorrect E says "No," and continues moving the rider until S gives the correct color. E then records this angle in the second place for that color, i.e., 50° under II in column "yellow." Make 2 or 3 trials at option of the instructor. The trials should be taken in haphazard order for each of the 4 primary colors. Record results as follows:

Color Zones of Right Horizontal Axis of Right Eye

	Yellow		*Red*		*Green*		*Blue*	
Trial	I	II	I	II	I	II	I	II
1								
2								
3								
Av.								

If a colorblind individual was found in Part 2, map out his color zones for the right horizontal axis of right eye as was done with normal subject.

Display and explain to the class an enlarged copy of the perimeter chart showing the color zones of the retina (opposite p. 188 in B. B. Breese's *Psychology*, 1921), or some similar chart.

Results and Discussion.

Part 1.

(1) Criticize the visual acuity tests and the astigmatism test, stating why it is better to use the E-test in many situations.

(2) Draw, label, and explain small diagrams for the following: emmetropic eye, myopic eye, hyperopic eye, lenses to correct the last two (see Averill for help).

Part 2.

(3) Compare a normal and a colorblind individual in their responses to colorblind tests. How important is normal color vision in everyday life? If an individual is colorblind is it important for him to know it? Why?

(4) Describe and criticize the colorblind tests used in this experiment.

Part 3.

(5) For the normal S obtain the average for the items listed under each of the four columns headed (I). Obtain the average angular distance from fixation at which each color was recognized. From these averages draw conclusions concerning the limits of the color zones of the retina (right horizontal axis).

Do the same for color zones of colorblind individual.

Compare color zones of a normal eye with those of a colorblind one.

(6) Draw and explain a schematic diagram showing the color zones of the eye (see Woodworth, p. 353). Try to account for the fact that we are not aware of the three color zones in everyday life.

(7) How does the Ladd-Franklin theory of color vision

account for red-green blindness (see Woodworth, pp. 353-354)?

There need be no general summary and conclusions for this experiment.

References

See end of experiment No. 11.

CHAPTER XIV

Experiment No. 11

SPECIAL PHENOMENA OF VISION *

II. *Blind Spot, Double Images, After Images, Contrast, and Color Mixture*

We are not ordinarily conscious of some phenomena of vision which nevertheless play a part in our daily life. Certain of these are easily brought to our attention by experimental means. The purpose of this experiment is to demonstrate the following special phenomena of visual sensation and perception: (1) the blind spot; (2) double images; (3) visual after-sensations; (4) brightness and color contrast; (5) color mixture.

Apparatus.—3″ x 5″ card with cross, dot and square; Brewster stereoscope with slides 10 and 38 (Titchener Series); 2 gray cardboards, one with red cross at center; a desk lamp; a black and a white cardboard with a 2-inch square of gray paper in the center of each, a black cardboard with a small square hole in its center; a color contrast frame; color wheel and color discs. (See Appendix.)

Method.—This is a lecture-demonstration experiment.† The instructor acts as experimenter and the class as subjects. The students are to take notes for use in their reports.

* For those classes which do not perform experiments on Color Mixture, Visual Adaptation and After Images, and Perception of Solidity and Distance.

† The instructor will find it necessary to expand the explanations.

Part 1. *The Blind Spot.*—The retina, which constitutes the innermost layer-membrane of the eyeball, is the specific sense organ for vision. It is complex and consists chiefly of nervous structure. In the center of each retina, opposite the pupil, is a slight depression which is called the fovea. Objects in direct vision are focused upon the fovea since it is the part of the retina yielding clearest vision. Somewhat to the nasal side of the fovea lies the blind spot, a small area (about 1.5 mm. in diameter) where the nerve fibres of the retina converge to form the optic nerve. In this region, known as the *optic disc,* there are no rods and cones, i.e., minute sensory end organs in the retina. Since no vision is possible there, the optic disc is commonly known as the *blind spot.* The existence of the blind spot and the way in which one eye supplies the deficiency caused by the blind spot of the other are demonstrated as follows:

To each student is given a card upon which is a black cross, dot, and square. To investigate the blind spot, close the left eye, hold the card with its long dimension horizontal and with the cross at the left. Starting at arm's length move the card *gradually closer,* always fixating or focusing the cross. It is absolutely essential to maintain fixation on the cross. Practice this a little if necessary. First, at a certain distance from the eye, the square should disappear. Maintain the card at this position for a moment and note whether a hole appears in the card where the square has disappeared, or whether the spot seems to be filled in with the background. Still maintaining the fixation of the cross, open the left eye and note that the square reappears. Now close the left eye again and continue moving the card slowly toward the eye. The square reappears and at a certain distance the dot disappears. As before, hold the card still, open

the other eye, and note the reappearance of the dot. Jot down notes concerning what happened during the experiment. Return cards to instructor.

Part 2. A. *Double Images.*—When images of two like objects, received separately by the two retinæ, are combined in such a manner that they are perceived as one object the images are said to fall on corresponding points of the two retinæ. However, when the images fall on non-corresponding or disparate points either single vision or double vision (double images) results. Single vision arising under these conditions, due to the fusion of two slightly disparate images, leads to perceiving the object as solid, i.e., having three dimensions. Double images are easily demonstrated. Fixation of the eyes on one point causes double images to arise from other points in the field of vision. The following is an example: Hold the index finger and a pencil in a vertical position at the level of the eyes, the finger about 2 feet and pencil 8 inches from the face. Be sure that the finger and pencil are lined up directly in front of your nose. Fixate steadily on the far object (finger) and at the same time note the appearance of the near object. Two images of the pencil should be seen. Hold the fixation and close one eye, then open it and close the other, noting the behavior of the two images. Similarly fixate the near object and note appearance of the far object. Maintain fixation and close right eye and left eye alternately. Which side, right or left, of the double image disappears when the near object is fixated and the right eye closed? When the far object is fixated and right eye closed? Fixate the near object and gradually move the far object (finger) toward the near one. How do the double images change with distance? (The instructor explains crossed and uncrossed images and their function in per-

ception of distance, see Pillsbury or Myers.) Each student now examines slide No. 10, 5 or some other simple figure of the Titchener Series in the Brewster stereoscope and notes the perception of solidity.*

Part 2. B. *Retinal Rivalry.*—Retinal rivalry is defined as the irregular alternation of colors or figures when the two eyes gaze upon different fields which do not fuse yielding unitary perception. Thus if a stereoscopic slide is prepared with colored squares, green for one eye and red for the other, and is then viewed through the stereoscope the colors alternate in continuous rivalry, first one and then the other color being perceived. This fluctuation of perception is due to shifting of the attention from one colored square to the other. This may be demonstrated with stereoscopic slide No. 38 of the Titchener Series. View blue with one eye and yellow with the other, or green with one and red with the other.

Part 3. *Visual After-Sensations.*—Visual after-sensations which are also called *after-images,* are visual sense impressions occurring after stimulation of the eye has ceased. The positive after-sensation is like the main sensation but weaker. When the after-sensation is opposite; as black after-sensation after a white stimulus, or red after a green stimulus, it is called a negative after-sensation. The main reason for after-sensations is the continuance of response of the sense organ after the stimulus is withdrawn. After the stimulus ceases, the sensation persists for a time; the positive after-sensation first, followed by the negative after-sensation. Demonstrate as follows:

Place a desk lamp containing a very bright light, 100-

* The floating-finger illusion may be demonstrated as an illusion produced by double images. See W. L. Sharp: *The Floating-Finger Illusion,* Psychol. Rev., 1928, Vol. 35, pp. 171-173.

150 watts, on the lecture table with reflector directed toward the class. Draw shades at windows to darken the room somewhat. Instruct each student to hold a book or another opaque object up before his eyes and rather close to the face for about ten seconds. Then, have him, while looking directly toward the light, lower and raise the book rather quickly several times, maintaining it in the lowered and raised positions, for about 2 sec. each trial. To perceive the positive after-sensation clearly, the book must be brought to the raised position, between the eyes and the light, with a quick jerk. When the book is in raised position note the bright spot of light which rapidly fades away. This is the positive after-sensation. A partially darkened room is desirable for achieving best results.

Present to the class the two gray cardboards, one with a red cross in the center. Time the students while they look steadily at the red cross for 25-30 seconds. At a signal from the instructor ask them to shift their gaze to the blank cardboard and note the after-sensation.

Part 4. *Brightness and Color Contrast.*—After viewing a dark surface, one of medium brightness looks light. On the other hand, looking from a light surface to one of the same medium brightness (used first) causes the latter to appear dark. In like manner, looking at one color and then at its complement causes the complement to look brighter. This is called *successive contrast.* Contrast effects arising from distribution of visual stimuli within the same field of vision is known as *simultaneous contrast* and is demonstrated as follows: Use the black and the white cardboards with a small square of gray at the center of each, and the black cardboard with a hole at its center. Allow the class to view the gray on black and on white, side by side, and note

the contrast. Now cover the white cardboard with the black one containing the hole. This will demonstrate that the grays are really equal in brightness. For color contrast show the contrast frame, asking the students to name the color of each section of the narrow horizontal band across the center. Lift up the tissue paper to demonstrate that the contrast effect is less pronounced when there are sharp boundary lines between the gray band and the colored surfaces. Contrast effect with colors is favored by proximity of the contrasting surfaces, indefinite boundary lines, lack of simultaneous light contrast (one surface more highly illuminated), and high saturation of inducing color. Contrast effects always operate in the direction of greatest opposition, i.e., white toward black, green toward red.

Part 5. *Color Mixture.*—By means of the color pyramid, color circle, the spectrum and other charts, the instructor explains to the class: complementary colors, hue, tint, chroma (saturation), the brightness series (grays) *vs.* the color series, and the "psychological" mixing of colors.

(A) *The Mixture of Complementaries.*—Using 2 complementary colors, as yellow and blue, demonstrate a mixture that is either yellowish or bluish-gray, i.e., is not a neutral gray. Readjust the discs so that a neutral gray is produced. Do the same for red and blue-green.

(B) *The Mixture of Non-Complementaries.*—Demonstrate the mixing of red and yellow to produce orange, and red and blue to produce violet.

Results and Discussion.

Part 1.

(1) Why don't you see a dark hole in the card at the place where the square or dot disappeared?

(2) Is the blind spot apt to be a hindrance to vision for the normal person? For an individual with vision in only one eye? Why?

Part 2. A.

(3) Summarize the results of the observation of double images.

Part 2. B.

(4) What are the chief descriptive characteristics of retinal rivalry, and under what conditions does retinal rivalry take place?

Part 3.

(5) Summarize the results of the after-sensation exercise and state the general law for the color of negative after-sensations. (See Gates, p. 156, or Woodworth, p. 355.)

Part 4.

(6) Summarize the exercise on contrast-effect. Name at least two practical applications of contrast-effects.

Part 5.

(7) State fully the principles of color mixture (mixing visual stimuli). How do the results of mixtures in this experiment differ from those of painters in mixing pigments? Give reasons for the difference.

(8) How does the Ladd-Franklin theory of color vision account for the general facts of color mixture? (See Woodworth, or other references.)

There need be no general summary and conclusions for this experiment.

References

For visual defects and visual tests, see L. A. Averill: *Educational Hygiene*, 1926, pp. 276-322; L. M. Terman: *The Hygiene of the School Child*, 1914, pp. 245-279.

For vision in general (color vision, double images, etc.), see R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 342-361, 397-402, 412-420; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 148-162; H. C. Warren: *Elements of Human Psychology,* 1922, pp. 57-84. J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 105-113; B. B. Breese: *Psychology,* 1921, pp. 166-196, 212-222; E. B. Titchener: *Text Book of Psychology,* 1911, pp. 50-92; C. S. Myers: *Text Book of Experimental Psychology,* 1911, Pt. I, pp. 71-101, 262-272; W. B. Pillsbury: *Fundamentals of Psychology,* 1922 (Rev. Ed.), pp. 109-135, 298-325.

CHAPTER XV

Experiment No. 12

IMAGINAL (IDEATIONAL) TYPES

Objects and situations not actually present may be represented mentally in the form of *images,* or, more technically, in the form of *ideas.* Images play an important part in the mental processes of "perceiving," "remembering," "imagining," and "thinking." Images appear in general to be less definite in form, and less stable and complete than sensations.

The most clearly demonstrable images are the positive and negative visual after-images of the preceding experiment. The images of dreams, and of hallucinations are only slightly less vivid and definite. Images which occur in imagination and in memory are usually less lifelike and complete, while the images which occur in "thinking" are commonly least vivid and most difficult to discover of all. It appears in general that much meaning may be represented by a very vague and fleeting image, or indeed merely by explicit or implicit motor reactions without the presence of discoverable imagery at all. Some persons of the "motor" and "verbal" types declare that all of their thinking progresses in terms of words and of actual kinæsthesis.

When several persons "think of" or "have an idea of" a dinner-bell, common-sense declares that they have the same idea or the same thought. Scientific observation, however, shows that while one person's idea may be primarily *visual* (he is said to have a "visual idea

of the bell''), another person's may consist primarily of characteristic auditory images; a third person may have a *kinæsthetic* image of ringing the bell; a fourth person may represent the bell to himself by actually making incipient movements with his hands, and thus perhaps arousing *kinæsthetic sensations* rather than true images; a fifth may merely say the word ''bell'' to himself (perhaps an *auditory image of the words as heard,* and also *kinæsthetic sensations* from inaudible movements of his throat as if speaking the word); others may find quite different combinations of representative imagery. Images with qualities corresponding to those of the other sense departments—cutaneous, taste, smell, and organic sensations—are much less often reported than those belonging to vision, audition and kinæsthesis.

As the above discussion indicates, individuals give evidence of considerable differences with regard to (1) their possession of imagery of the various kinds, as well as with regard to (2) the ease with which given images appear and the stability with which they continue, (3) their vividness, (4) the completeness or accuracy of their images, (5) the frequency with which a given sort of imagery is used, and (6) the degree of versatility shown in shifting from one sort of imagery to another. Thus mere possession of vivid imagery of a given sort need not necessarily imply that such imagery will be accurate, or will appear frequently. Furthermore, and quite contrary to the common notion, few individuals are limited to the possession or use of imagery of any single sort. Few if any persons belong to the ''pure visual,'' ''pure auditory,'' or ''pure kinæsthetic,'' or ''pure motor'' type. Many good visualizers, for example, turn to auditory and kinæsthetic imagery in many situations where words are concerned. Such persons

are said to be "visual" for concrete objects and "auditory-kinæsthetic" in dealing with more abstract verbal material. Imaginal type, that is to say, is not an entirely fixed, definite, and permanent characteristic of an individual, implying almost complete dependence upon one variety of imagery. It is at most a tendency to favor one or several sorts of imagery, and may shift with age, with problem, with purpose, with situation, and with a host of other conditions.

The purpose of this experiment is: (1) to study and describe the following kinds of imagery: visual, auditory, kinæsthetic, tactual, olfactory, gustatory, thermal, and pain; and (2) to determine whether there is any marked tendency for one or several forms of imagery to predominate.

Apparatus.—Group of pictures for report from memory, 9 letter-squares, timer. Problem questions. (For details, see Appendix and *Notes for Instructors.*)

Method.—The present experiment is intended to bring about a wide variety of situations as favorable as possible for the observation, now of one, now of another sort of imagery. In every case, the introspective observation which constitutes the basis of report is difficult. It is especially so for unpracticed subjects. Two special warnings must be regarded if you are to make adequate and reliable reports.

Do not confuse actual imagery with the particular form of imagery which you may have most reason to expect from the nature of the conditions or of the problem. Although, for example, you may be able to report the objects in a picture *visually* presented, with their correct coloring and arrangement, this ability to report accurately does not prove that you have used visual imagery. You may have depended mainly on a

rapid *auditory-kinæsthetic* naming of objects and a general verbal description of the picture, or on some other substitute for visual images. If, again, you try to call up the odor of a rose, you may think that you "have" it in olfactory terms, but discover by more careful observation that what you are getting is very different in character. You may find (1) a visual idea of a rose, perhaps of a certain definite color and variety, (2) a tendency to sniff (as in smelling) and to assume a facial expression indicative of pleasure, (3) a feeling of moving your hand to bring the rose nearer your face, or the visual image of such a movement, (4) certain descriptive words, such as "rose," "fragrant," etc. With no olfactory imagery at all you may very likely report that you have images of odors unless you are very careful.

Try to avoid confusion of present sensory processes with imagery. For example, when you seem to feel yourself humming a familiar tune, try to discover whether the suggested movements are wholly imaginary, or whether you are actually making certain of the movements, even though they may be reduced in amount. And when you imagine a movement of your arm, be on the watch to see whether true kinæsthetic images or merely slight actual kinæsthetic sensations are present.

(a) The instructor presents visually a group of pictures for 30 seconds. The class is required to observe, and later to report as completely as possible the objects presented, the colors, sizes, etc. Each student should record (1) the description above called for, (2) an account of the method used in recall, with special reference to imagery or actual sensory processes, (3) an account of his procedure while looking at the pictures, with special references to the use of imagery and to the presence of tendencies and feelings of movement.

(b) Each member of the class prepares in advance a number of diagrams according to the instructor's directions. The instructor then presents visually a series of "letter-squares," for about 15 seconds each. Each student is to observe, with intention to report as completely and accurately as possible, the letters, numbers, designs, etc., in the squares. In certain cases special instructions as to mode and order of recall will be given.

Immediately after the card is removed, members of the class enter in their diagrams as many letters, numbers, or other symbols as they can recall. Each S summarizes briefly his method of recall, noting especially (1) whether or not there appears to be any imagery present; (2) whether there is any *visual* imagery, and if so how much and stable this is; (3) whether there is a tendency to *say over* the letters—either kinæsthetic imagery or actual movement; (4) whether there is a tendency to hear the letters as spoken (auditory imagery); (5) whether eye-movements appear during recall. Each S later checks all errors and enters the correct symbol beside the wrong one for later reference.

(c) A simple problem is presented, the statement of which has been written on the blackboard and covered from sight. S solves the problem as rapidly as possible, noting the answers as (1), (2), (3), and (4) and recording as (5) the time required for the solution, as read from the timer which is started on presentation of the problem. A brief introspective report is also to be recorded.

(d) If time permits, the instructor reads to the class certain additional problems, calling for solution and introspective comment.

(e) The preceding exercises have been concerned with the imagery which appeared in connection with certain

particular situations when the attention was primarily on the task to be done and not on the arousal of any particular images. In the following exercises of the questionary, S makes a definite attempt to summon images of different kinds, and estimates the degree of *vividness* of the imagery which appears. Each image is graded on the following scale: 0 = no image; 1 = faint; 2 = distinct; 3 = very vivid. S also adds any introspective comment that he can offer regarding the characteristics and peculiarities of his images, and notes especially the presence and character of any indirect secondary or "unexpected" imagery.

I. *Visual.*—Indicate the degree of vividness with which you can image each of the following *in visual terms:* (1) the *color* of an American Beauty rose, (2) its *form,* (3) the *brightness* of a white cup, (4) the *brightness* of snow in the sunshine, (5) *a group of colors* in a bunch of sweet peas, (6) the face of a person whom you know well, (7) this person laughing or making a characteristic gesture, (8) your name in your own handwriting, (9) the name of this university printed.

II. *Auditory.*—Indicate the degree of vividness with which you can image each of the following *in auditory terms:* (1) the report of a gun, (2) the crunching of snow under your feet, (3) the characteristic tone quality of a violin, (4) the tone quality of a cornet, (5) the air of "Yankee Doodle," (6) a selection as played by a whole orchestra, or sung by a quartet, (7) the words of a familiar Mother Goose rhyme, (8) the words of these instructions as you read them.

III. *Kinæsthetic.*—How vividly can you image in *kinæsthetic terms:* (1) rocking in a chair, (2) biting a lump of sugar, (3) clenching your fist, (4) singing the air of "Yankee Doodle," (5) a facial expression of fear,

(6) the bleating of sheep, (7) singing a very high tone?

IV. *Tactual.*—With what vividness can you form a *tactual image* of (1) the feeling of velvet, (2) of smooth glass, (3) of sandpaper, (4) of mud, (5) of the impression of a dime on the palm of your hand, (6) of a line traced by the point of a pencil on the back of your hand?

V. *Olfactory.*—How vividly can you image *in olfactory terms* the odor of (1) coffee, (2) camphor, (3) onions, (4) apple-blossoms, (5) decaying fish, (6) kerosene, (7) carnations?

VI. *Gustatory.*—How vividly can you image the following *taste qualities:* (1) sweet, (2) salt, (3) sour, (4) bitter; the so-called "taste" of (5) an apple, (6) an olive, (7) a beefsteak?

VII. *Thermal.*—How vividly can you image (1) the sensation from stimulation of a cold spot, (2) ice cream on the tongue, (3) a draught of cold air, (4) the sensation from the stimulation of a warm spot, (5) a warm bath, (6) the warmth from a fireplace?

VIII. *Pain.*—How vividly can you secure an image of the pain of (1) the prick of a pin, (2) running your finger along the edge of a sharp knife, (3) toothache, (4) the stimulation of a pain spot, (5) the prickly pain that you get when your foot has gone to sleep?

Results and Discussion.

(1) Can you discover any differences in the *vividness* of images corresponding to the various sensory qualities as these appeared under the conditions presented by the questionary? To answer this question find the *median* value assigned to the images which appeared under each one of the headings I-VIII of the questionary of exercise (e).

(2) When you were filling out the questionary did you

find any particular kinds of imagery appearing persistently when you were looking for other forms? Specify.

(3) In exercise (b) do any of the errors appear to be indicative of the imagery used, i.e., due to confusion of like-sounding, or of like-appearing letters or symbols?

(4) Considering the results of exercises (a)-(e) what statements can you make regarding your imaginal type? Can you classify yourself as belonging to any one type? If not, what tendencies do you find with regard to predominance of, or special dependence upon, one or several forms of imagery, or upon motor reaction?

(5) What have experimental results shown concerning imagery types? (See Starch reference below.)

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 109-116; E. B. Titchener: *Text Book of Psychology*, 1911, pp. 194-200, 364-370, 416-421, 505-521; J. B. Watson: *Behavior,* 1914, esp. pp. 16-21; K. Dunlap: *Elements of Scientific Psychology,* 1922, esp. pp. 298-304; M. R. Fernald: *Psychological Monographs,* No. 58, 1912; W. B. Pillsbury: *Fundamentals of Psychology*, 1922 (Rev. Ed.), pp. 258-263; D. Starch: *Educational Psychology,* 1926 (Rev. Ed.), pp. 181-184.

CHAPTER XVI

Experiment No. 13

THE LEARNING CURVE *

Ordinarily improvement in learning is very irregular. The rate of progress varies widely for different subjects and also from time to time for the same individual. The rate of improvement is sometimes rapid, sometimes slow. Increase in efficiency of performance due to practice is shown by the *learning curve* or *practice curve.* The general form of such a curve may be influenced by the nature of the function involved in the learning; by the ability, individual habits, and previous experience of the subject; and by the environmental conditions under which the task is performed.

For simple, definite tasks the limit of improvement through practice is soon reached. This limit at which a habit cannot be supplanted by a more efficient one is termed the *physiological limit.* With more complex performances improvement is possible over a relatively long period of time and with very complex tasks it is practically interminable. In everyday life few people come anywhere near their possible limits of efficiency.

In this experiment we shall study the increase in effi-

* Two important studies constitute the classical foundations of learning as a subject of quantitative and objective psychological investigation: (1) Ebbinghaus' monograph on the learning and recall of nonsense syllables published in 1885; and (2) Bryan and Harter's study of the practice curve in learning telegraphy (1897). The first practice curve was, however, constructed in 1886-1887. Following these early experiments the study of learning has continued steadily up to the present time.

ciency due to practicing a relatively simple performance. The problem is: (1) to obtain a learning curve in a letter-digit substitution test; (2) to note the general form of the curve; and (3) to study the uniformity in rate of improvement during the learning.

Apparatus.—Timer and mimeographed material (see *Notes for Instructors*).

Method.—The instructor acts as experimenter and the class as subjects. After cautioning the students to keep the blanks face down until otherwise instructed, a set of 20 mimeographed sheets are distributed to each subject. Read the following directions: "*This is a letter-digit substitution test. At the top of the individual experimental sheets is a series of five letters with a number underneath each one.*" (Illustrate on the blackboard but do not use the letters on the blank.) "*Following this key are several lines of letters. Your task is to begin with the first line and place underneath each letter the digit corresponding to it according to the key. Proceed as rapidly as possible but be certain that your work is accurate. When I say 'ready,' 'go,' turn over the top sheet and begin. Continue working until I say 'stop.'*"

Be sure that everyone understands what is to be done and then give the 20 practice trials, one on each sheet. Allow 1½ minutes per trial and about 1½ minutes of rest between trials. During the rest period have each student write at the top of the page just finished the serial number of that trial and his score (number of letters coded) and then immediately place the sheet at the bottom of the pile.

When the 20 trials have been completed, have the members of the class read off their scores while the instructor adds them on a machine. All scores for trial

one are taken first. Place the average scores for each trial of the class data on the blackboard for the students to copy. (Medians rather than averages may be used if a calculating machine is not available. Curves plotted from averages and from medians are practically identical.) Data for the experiment will consist of the subject's scores and the class averages (or medians).

All students are to write up this experiment.

Results and Discussion.

(1) Make out a table showing the individual and the class (averages or medians) results for each trial. Draw learning curves for S and for the class on the same chart. What is the general form of these learning curves? Are the curves typical? Why?

(2) Compare the individual curve with the class curve. Which shows the greater "smoothness" or "regularity?" Explain why this is so. Also explain any other marked differences discovered between the two curves.

(3) Smooth the class data by means of the moving average and record the results in a new table. Use these smoothed data and on the graph called for in result (1) and with the same base line and scale, plot the learning curve. Explain in detail the difference between this curve and the class curve plotted from unsmoothed data. Of what use is smoothed data?

(4) Obtain for the class data the differences between first and final score in each quarter and in each half of the practice series by subtracting first from fifth score, sixth from tenth, etc. Record this material in a table. By reference to the table point out periods of fastest and of slowest learning. Is there

any final spurt evident in the last one or two trials in either your own or the class curve? If so, suggest a possible reason for the spurt.

(5) What are "plateaus" in learning curves? How are they explained? (See Pyle, p. 32; Jordan, p. 77; Sandiford, pp. 213-217; Starch, p. 164; or Gates, p. 326.) Did you find any plateaus in your curves?

(6) Would you expect all learning curves to be of the form found in this experiment? Why? (Consult references if necessary, especially Starch, pp. 154-164.)

Summary and Conclusions.

References

R. Pintner: *Educational Psychology*, 1929, pp. 197-221; W. H. Pyle: *The Psychology of Learning*, 1928 (Rev. Ed.), pp. 17-38; D. Starch: *Educational Psychology*, 1927 (Rev. Ed.), pp. 154-190, esp. pp. 154-167; A. M. Jordan: *Educational Psychology*, 1928, pp. 72-79; P. Sandiford: *Educational Psychology*, 1928, pp. 207-236, esp. pp. 207-220; A. I. Gates: *Elementary Psychology*, 1928 (Rev. Ed.), pp. 320-327; R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 124-151; F. A. C. Perrin and D. B. Klein: *Psychology*, 1927, pp. 260-268; W. F. Book: *Learning to Typewrite*, 1925, pp. 269-288.

CHAPTER XVII

GRAPHIC REPRESENTATION OF THE LEARNING CURVE

The Learning Curve.—One of the simplest and most common forms of graph used in psychology is that which represents the progress and other characteristics of learning and practice.

Suppose that two previously unpracticed subjects, A and B begin learning to write words in the Braille alphabet (raising with a stylus the dot-point symbols read by the blind). They practice for fifteen minutes every day, and their progress is measured by the number of letters correctly coded in that length of time. Table 9 shows the results for twenty consecutive days.

Inspection of the figures in the table shows that A is a more rapid and better learner of Braille than B, but many other points of comparison and variety of characteristics of learning in general appear readily only if we represent the data graphically.

About two inches from the bottom of a sheet of cross-section paper draw a horizontal base line, the abscissa scale or axis of *abscissæ*. From left to right mark off on this line equal divisions to represent the successive periods of learning (in this case, the days from one to twenty). Near the left end of this base line draw a vertical line, the ordinate scale or axis of *ordinates*. Make this line into a scale by laying off on it, from below upwards, sub-division points to represent the measures of degree of learning (in this case, the number of letters coded in fifteen minutes). Again be sure to choose a

size of unit which just allows the scale to cover the range of values to be represented without exceeding the dimen-

DAY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	136	188	264	307	352	374	391	415	427	420	443	456	460	468	459	465	480	484	482	485
B	107	149	188	236	270	314	332	335	350	354	387	368	372	390	385	396	422	428	433	436

Table 9. Progress in learning Braille by two subjects, as measured by symbols coded on successive days.

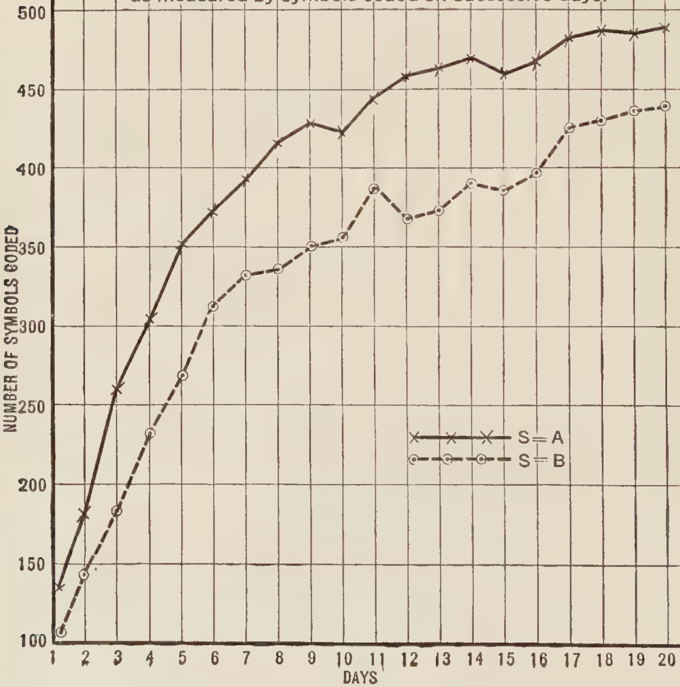


Fig. 14. Learning curves for subjects A and B, representing in graphic form the data of Table 9.

sions of the plotting sheet. Here, for example, the largest value is 485; the smallest 105; the range therefore 380; and a convenient division unit 10. Numbers should be placed on both scales exactly *opposite* the

points they indicate, and *not between* two points. Bear in mind that in graphic representation both the base line and the vertical are *scales,* and that *points* on them should accordingly be labeled very precisely. Label each axis to indicate exactly what it represents, number and entitle the table and graph as a whole.

Now make tiny X's or large dots exactly above each point on the base line at distances (measured by the ordinate scale) equal to the corresponding successive measures of A's learning. Then connect these X's or dots by straight lines, thus "plotting" A's learning "curve." Above the same base line plot a similar curve for B's results.

This graph (Fig. 14) not only represents exactly all the data of our original table, but it also calls to our notice a variety of further generalizations; such for example as that A surpasses B quite constantly; that efficiency in learning as measured by letters coded increases rapidly at first and then much more slowly; that the progress of an individual's learning is not quite regular, but subject to occasional spurts, slumps, and halts (indicated by humps, hollows, or plateaus on the curve). The latter observation might lead us to try to discover unusual factors or conditions which may have been responsible for such results as those of A on the tenth day, and of B on the eleventh, etc. Other uses of learning curves, and the use of similar curves for representing preferences and the like, will appear in later experimentation.

It is worth noting here that the progress of learning may be represented not only by work done, but by the time required, or by errors made in doing a given amount of work. In such cases a descending rather than an ascending series of measurements denotes progress. The

curve therefore is of the reverse type, and takes the form illustrated in Figure 15.

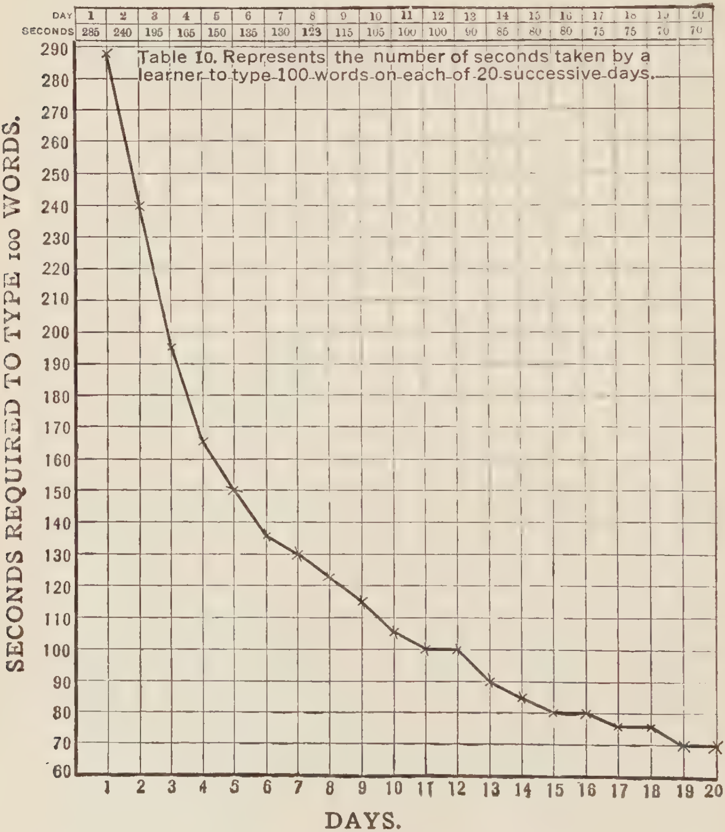


Fig. 15. Learning curve for typewriting, representing data of Table 10.

The "Typical" Learning Curve.—Several factors such as nature of the function measured, ability of the subject, etc., affect the form of the learning curve.

However, a rather common form taken by the practice curve in learning many complex tasks is termed "typical learning curve" by many writers. Such a curve rises (drops if time or error scores are employed) rapidly during the first part of the learning, and then more and more slowly as the learning progresses. These curves frequently have "plateaus"; a rather flat stretch (period of little improvement) in the practice curve followed by a second rise. "Typical curves" are illustrated on page 321 of Woodworth's *Psychology*, on page 73 of Jordan's *Educational Psychology,* and page 322 of Gates' *Elementary Psychology.*

"Smoothing" the Learning Curve.—Customarily there are fluctuations in the rise (drop for time or error scores) of a practice curve. On certain trials a combination of several favorable circumstances may cause a rapid rise in the learning curve; or a combination of several adverse ones bring about a sudden drop since many factors influence learning. Thus originate the fluctuations noticeable in most learning curves. If the curve is "smoothed" we perceive more easily the general tendency of the progress in learning. Several methods are employed for smoothing data. In the frequency polygon the method of grouping was used. The moving average is a handy method to employ with learning curves. The procedure is as follows: Multiply the record of trial one by the number two, add the second record, and divide by three; this yields the first score of the smoothed curve; obtain the average of the first, second, and third records for the second score of smoothed curve; continue in this manner, taking the average of three scores each time. Double the final record and combine this with the next to the last score, and then divide by three for the final score of the smoothed data.

For example, a subject made the following numbers of errors in ten successive trials in a simple maze (maze incompletely learned):

	Successive Trials									
Actual data	34	10	20	14	9	7	11	6	5	5
Smoothed data ..	26	21.3	14.6	14.3	10	9	8	7.3	5.3	5

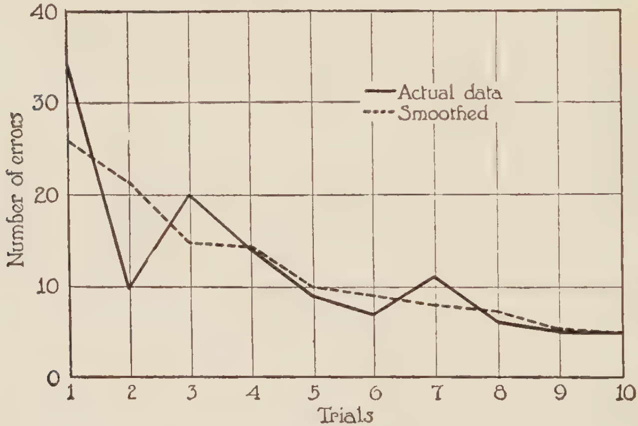


Fig. 16. Original and smoothed learning curve.

The second row of figures represent the scores in the first row smoothed by the moving average method. These data are plotted in Figure 16. Note the contrast in the form of the two curves.

Exercises.

(1) In learning to run a maze blindfolded an S made the following number of errors in 24 successive trials: 74, 31, 42, 27, 8, 11, 5, 12, 19, 8, 15, 11, 7, 4, 1, 1, 4, 0, 3, 1, 2, 0, 0, 1. Represent the successive trials by equally spaced divisions on the base line. Represent the successive numbers of errors as ordinates, and draw the error learning curve.

(2) Another S ran the maze in successive trials in the following numbers of seconds: 193, 105, 85, 82, 55, 42, 50, 37, 24, 19, 34, 17, 19, 18, 18, 22, 17, 15, 24, 21, 17, 15, 13, 14. Plot his time learning curve. Smooth the data and plot the smoothed curve on the same graph.

(3) A child made the following numbers of errors in successive trials in a simpler maze: 35, 9, 18, 11, 9, 13, 6, 6, 5, 5, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0. Plot his learning curve.

(4) In learning Braille (blind alphabet) an S coded correctly in 15 minutes, on each of 20 successive days, the following numbers of words: 1st day, 189; 2nd, 198; 3d, 236; 4th, 262; 5th, 324; 6th, 342; 7th, 384; 8th, 403; 9th, 401; 10th, 443; 11th, 427; 12th, 448; 13th, 419; 14th, 450; 15th, 499; 16th, 497; 17th, 515; 18th, 514; 19th, 538; 20th, 552. Represent the data in a learning curve.

References

H. E. Garrett: *Statistics in Psychology and Education*, 1926; pp. 71-74; W. H. Pyle: *The Psychology of Learning*, 1928 (Rev. Ed.), pp. 17-38; R. Pintner: *Educational Psychology*, 1929, pp. 197-221.

CHAPTER XVIII

Experiment No. 14

HABIT FORMATION IN THE MAZE

We have said that from the point of view of behavior the animal begins life as a "bundle" of reflexes. The word "bundle" was there used to indicate that in most animals these reflexes function, not in entire independence of one another, but (owing to a central nervous system and to its inherited connections) with some degree of coördination and mutual influence. Through the special "openness" of certain nervous pathways, certain reflex arcs may be so coördinated at birth as to produce tendencies or "sets" toward relatively definite and ordered *series* of responses, whenever the appropriate stimulus or situation provokes the initial preparatory reaction of the series. Such inherited "sets" are called *instincts.*

In addition to their native reflex and instinctive tendencies, however, all animals possess in a greater or less degree the capacity for *learning*, that is to say, for *modification of behavior through experience.* Upon the possibility of such modification, and upon the fixity with which the modifications persist, depends what in general we call the formation of habits.

The capacity for forming useful habits depends upon a variety of factors:

1. *The adequacy of sensory equipment and of the environment which affects the senses.* The amœba cannot learn to read, in part, for lack of eyes; the deaf have

difficulty in learning to talk; man fails to use smell to as great an extent as the dog because the human nose is far above the ground and out of the region where smells cling.

2. *The original stock of reflex pathways and of instinctive interests and incentives.* Perhaps reptiles learn especially slowly, because, among other reasons, their incentives to action are fewer, or because they instinctively attend to relatively fewer features of the environment.

3. *The ease or difficulty with which the original reflexes and instincts may be modified.* Some animals at birth behave practically as do adults. Because their original reflex and instinctive arcs are so *fixed* (closely organized) and relatively perfect at birth, they can scarcely be said to have a period of infancy. For each situation significantly different for them, they appear to have a ready-made and fully organized instinctive response. This response may persist even under quite different (and therefore perhaps inappropriate) circumstances. Speaking roughly, the insects and crustacea are examples of this class. Man's nervous system, on the other hand, is extremely *plastic* or modifiable. His natural reactions are "loosely organized" and of great complexity and variety. They are in consequence easily modifiable by habit. His period of infancy is especially long, and his learning ability is correspondingly great.

4. *The permanence or "retention" of the modifications once they are made.* In certain experiments small-brained rats ultimately learned to run a maze as well as normal animals. Two months later, however, when they were required to relearn it, they needed about twice as much retraining as the normals, who had had less original practice.

5. *The adequacy of motor equipment.* Many animals appear more stupid than their place in the evolutionary scale would seem to warrant, because of clumsy motor organs. Possibly horses, birds, and reptiles may be taken as examples. On the other hand, raccoons and monkeys learn rapidly to carry out a great variety of unusual activities involving their prehensile paws. The flexible and complex hand and throat of man, together with the correlated manipulative and language abilities and the institutions and other forms of social heredity thus made possible are probably responsible in a great measure for his intellectual advancement over the other animals.

Because of the original loose coördination of their reflex and instinctive tendencies, unpracticed animals respond to many situations by a complex variety of little unified or "random" movements. The movements, that is to say, are not clearly or precisely suited to adjust the organism to the particular situation which calls them forth. Certain of the movements, therefore, appear superfluous or "useless." Young animals show this random activity most clearly, but even the human adult, in learning to solve a new or difficult problem, does learn in some measure by "trial and error" responses, either of the "explicit" sort made in solving a mechanical puzzle or of the "implicit" (hidden) sort which we call thinking or learning by observation.

The learning of lower animals is most conveniently studied through use of the problem-box (puzzle-box) or of the maze. With the maze, for example, the animal is put in a home-box at one end with food at the other outlet. A strong incentive to action is commonly furnished by hunger and the smell of food. The first successful run is made in a relatively *long time* and with a

great number of "errors" ("useless movements"). In later runs the time and the number of errors *gradually* diminish. In the end a habit is formed, in which the necessary series of acts is performed with relatively few superfluous movements.

In the case of man, incentives are more complex, are controlled with greater difficulty, and conditions for simple "trial and error" are harder to obtain. In this experiment the desire to make a good record and a consequent good social impression must furnish the chief incentive. Most of the unit actions involved (holding the pencil, moving in a straight line, etc.) are complex, coördinated, already learned movements, and not true reflexes. But the fact that S is blindfolded and can have at first no "idea" of the proper directions of movement practically compels him in the early stages of maze-learning to adopt a policy of random movements and trial and error learning.

In this experiment we shall make a quantitative investigation of habit formation in the maze and determine, as far as possible, the methods employed by a human subject in maze learning.

Apparatus.—Fiberboard maze "A" or "B," with cover; table clamp; stylus or blunt-pointed pencil; watch or timer; blindfold. (Ten-cent goggles with cardboard disks over inside of glass are suitable. A sheet of tissue paper [Kleenex] may be used for cleanliness.)

Method.—Fasten the maze firmly near the edge of the table by a clamp which grips it by the border near the starting point.

E is not to look at mazes other than the one he is actually to use! Under no circumstances should he let S see any maze! He should blindfold S or make sure S has his eyes closed before uncovering the maze. Later,

between trials, he must always be sure that the maze is covered before S opens his eyes or removes the blindfold.

E first studies the maze until he knows the "true path," and until he is able to tell with certainty what constitutes an error. Blind alleys are outlined in black.

An *error* is any move *which leads into a blind alley,* or which goes a perceptible distance *backwards in the true path.* When blind alleys are themselves crooked, the entrance counts as one error and each successive turn thereafter which leads away from the true path counts one more. Otherwise expressed, *each turning of a corner, or any other definite movement away from the true path,* is an error.

The safest way to record errors, especially at first, is by making on paper one short stroke for each error, and by finally counting these strokes. At later stages when errors become less numerous, every five marks may be recorded thus ~~////~~ to facilitate counting. Still later, simple counting without marks may suffice. Make sure that you can count errors by getting an instructor to try you out.

Read S the following instructions, *word for word:* "*Here is the starting box.*" (Place his pencil in it.) "*You are to slide your pencil along in the grooves until I tell you 'You're out.' Keep the pencil in the grooves and try both to reach the goal as quickly as possible and to learn to avoid wrong turns. Start when I say 'Go.'*"

Give no further explanations. Make no comments, criticisms or suggestions. See that S does not lift his pencil out of the grooves. Do not let him use his other hand to feel the maze ahead of his pencil. Both E and S must take the experiment seriously if they are to ob-

tain reliable measures. Give S no information about his own or his neighbors' results until the entire laboratory period is over. In particular, do not tell him the location of the goal, but let him find it himself in Trial No. 1.

Give the word to S. Count his errors and record in seconds the total time elapsing between "Go" and "You're out." Record in a table as "Trial 1." Then record in quotation marks all S's significant volunteered remarks.

Then try in a few minutes to find out and record:

(1) S's *feelings and attitudes.* (Was he generally discouraged? Did he feel "blocked" or utterly helpless at times? Did he stop working hard at such times? What evidence that he was working seriously or not seriously? Does he tend to go slowly and cautiously and to learn details bit by bit, or to go exploring rapidly and to take the maze as a whole, etc.?)

(2) *The means whereby he aimed, consciously, to meet the difficulties.* (How did he go about it? Was it mere "trial and error," or did he have a plan in advance, or make up little plans as he went? What plans or tricks? In what special way was he better off in Trial 2 and later than in Trial 1? etc.)

(3) *The direction of S's attention.* (Did he pay attention chiefly to his ideas about the maze as he went, or to the "feel" of his pencil and the "feel" of his moving arm? Did he talk to himself in inner speech? Take mental notes about it? If so, what notes? What sort of pictures? etc.)

Then take rapidly 9 more trials, with only brief rests intervening between them. Record only errors, time and important volunteered remarks. The successive times

may be somewhat irregular. Do not tell S, but simply try yourself to account under "Remarks" for especially large or small times. After Trial 10, get information as to changes regarding points (1), (2) and (3) above, remove blindfold and ask S to draw as accurate a diagram of the maze as he can. Include this drawing in your notes.

Keep on with similar trials and with similar records until S has made at least 25 trials. Secure a second drawing and a second report like the one taken after Trial 10.

DO NOT LET S SEE THE MAZE EVEN YET, HOWEVER, but if time permits, continue until S has given *three successive trials without error,* and until his time in one of those three trials is not greater than 15 seconds.

Results and Discussion.

(1) Briefly summarize your notes on S's verbal reports and behavior. Try to generalize (a) as to the changes in attitude, in behavior, and in consciousness and direction of attention as learning progressed, (b) as to whether S emphasized speed or accuracy in general, or at any particular stages of learning.

(2) To complete three errorless trials the median number of trials obtained from groups of 125 students each were: for Maze A, 36 trials; for Maze B, 39 trials. The typical median of the times for the first 25 trials (either maze) is 40 seconds; of errors, 8 errors. Compare the results of your subject with these medians and explain.

(3) Plot a learning curve for your S using all of his successive times as ordinates. The following *average times* of succeeding trials were obtained from

58 subjects who used Maze A and from 65 subjects who used Maze B:

Trial No.	1	2	3	4	5	6	7	8
Maze A	297	177	164	106	79	77	68	55
Maze B	426	255	151	140	91	80	70	74

Trial No.	9	10	11	12	13	14	15	16
Maze A	46	45	45	44	41	40	36	35
Maze B	63	58	60	55	49	40	36	38

Trial No.	17	18	19	20	21	22	23	24	25
Maze A	34	31	30	29	31	26	26	25	23
Maze B	35	34	33	30	32	30	30	28	25

Use appropriate values from the above table and on the same graph with S's time curve, plot as a dotted line an average time curve for the Maze (A or B) used by your S. Try to explain irregularities in S's curve by reference to your notes. Compare individual with average curve and try to explain differences discovered.

(4) Plot a second learning curve for S using *errors* as ordinates. The following *average numbers of errors* appeared from the results obtained by the same subjects (as in problem 3).

Trial No.	1	2	3	4	5	6	7	8
Maze A	74	43	49	30	22	24	20	15
Maze B	127	63	45	45	26	22	20	19

Trial No.	9	10	11	12	13	14	15	16
Maze A	12	11	10	12	9	8	8	7
Maze B	18	13	13	12	10	7	6	7

Trial No.	17	18	19	20	21	22	23	24	25
Maze A	7	6	6	5	6	5	4	4	3
Maze B	6	6	6	5	6	5	5	5	4

Use appropriate figures and plot the average error curve as a dotted line on the same graph as the error curve for S. Try to account for irregularities in individual curve. Compare individual curve with average curve and try to explain differences discovered.

Explain why the average curves in (3) and (4) show greater smoothness or regularity than the individual curves.

(5) State whether your S tends to stress speed or to stress accuracy (avoidance of errors): judging from S's verbal reports and volunteered remarks, and from your observations of his behavior; judging from a comparison of his results in (2) with the typical results; judging from the comparison of curves in (3) and (4). Taking all these judgments into consideration together, should you say that S clearly tends to favor speed or favor accuracy, or stresses the two about as an "average" subject does? Give reasons for your statement.

(6) Give a theory to account for the fact that "useless" movements tend to drop out (time and errors tend to decrease) as practice goes on, instead of persist-

ing by habit. (See Woodworth, pp. 166-168; Watson, pp. 251-276; Dashiell, pp. 331-348.)

(7) Summarize Perrin and Klein's description of maze learning with humans and compare their results (description of subject's behavior in maze and S's verbal reports) with those in this experiment. (See Perrin and Klein, pp. 244-248.)

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 131-136; J. B. Watson: *Behavior,* 1914, pp. 251-276; F. A. C. Perrin and D. B. Klein: *Psychology,* 1926, pp. 196-254; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 325-347; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 280-320; J. B. Watson: *Psychology from the Standpoint of a Behaviorist,* 1919, pp. 269-309.

CHAPTER XIX

Experiment No. 15

LEARNING AND TRANSFER OF TRAINING IN MIRROR DRAWING

Under normal conditions tracing with a pencil along a path in any given direction is performed with no hesitation or difficulty. We say that we have already formed habits which enable us to move the pencil in the appropriate ways and that such movements are guided or directed by sight. More technically, the visual and kinæsthetic impressions from the paper and lines, the pencil, the hand and so on, together with the auditory perception of the words "Trace so and so" are "cues" or stimuli which establish the appropriate set and "touch off" the appropriate movements.

If, however, the hand and paper are hidden from direct vision by a screen and the subject sees them only indirectly in a mirror, the cues to movement are greatly altered. Since the rays of light have undergone reflection before reaching the eye, an apparent conversion of directions in the field of vision takes place. The subject is compelled to interpret differently (react differently to) what he sees. In order to trace the path accurately, a new habit of coöperation between hand and eye must be formed. The formation of this new habit is especially difficult because it is opposed by the ingrained and persistently practiced habits of a lifetime.

Besides the light it throws upon the process of forming a new habit to replace temporarily an old one, the

experiment in mirror drawing has two other advantages. The learning process it involves runs at a relatively elementary level—there is little chance for the thinking, the scarcely realized adoption of tricks, or rational schemes which are so likely to enter into many problems of human learning and to result in sudden spurts which are not truly characteristic of the simple learning curve. That learning may (and often does) thus proceed by "trial and error" without our being explicitly conscious of its details or methods and without our specific intention or rational guidance, is a fact worth noting.

Mirror-drawing is especially suited also to bring out the phenomenon of "transfer" of practice. Suppose that we measure first the subject's initial ability with his more and with his less skilful hand; then give him practice with the more skilful hand until his ability with that hand is considerably improved; and finally measure again the ability of the less skilful hand. If this final measurement shows that the ability of the unpracticed hand is now much greater than before, we say that the gain of practice has "transferred" from the one hand to the other. Such an outcome would indicate that even such apparently simple learning is not in reality a simple matter of "training of the eye and of certain muscles," but that it involves the establishment of a general "pattern" of coördination, probably through complex neurological changes involving both sides of the brain.

The experiment demonstrates another point of considerable psychological importance, namely that what we take as a matter of course and call "seeing that objects are right or left, or near or far" is not an altogether simple and natural affair, nor indeed a mere

matter of "seeing," but rather a highly developed process of *perceiving,* made possible by previous trial and error *reactions* through a long period in childhood, and still involving implicit or explicit reactions toward the object seen.

The purpose of this experiment is to investigate trial

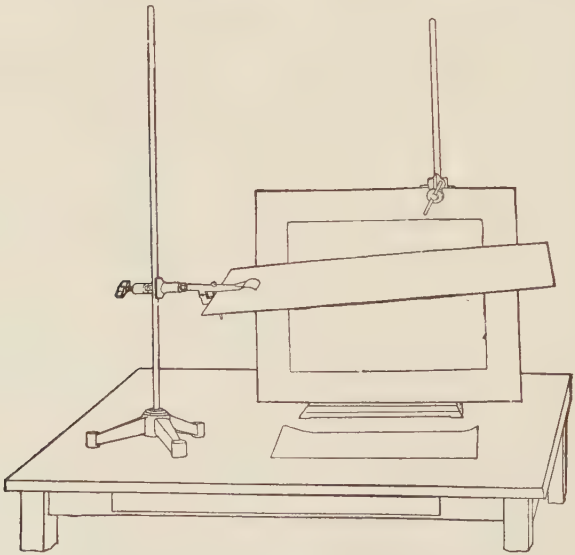


Fig. 17. Arrangement of apparatus for mirror drawing.

and error habit formation in the mirror drawing experiment; to determine the transfer of practice effect from skilled to unskilled hand; and to note how habits of perception arise through trial and error reactions.

Apparatus.—Mirror and support. Screen and support. Star patterns, pencil, timer. (See Appendix for further description.) 44 or

Method.—Make out a 54-row table with three narrow and one broad column. Entitle the columns: Trial No.;

time; errors; notes and remarks. Fill in the first column as follows: P1, P2 (for preliminary); Sup 1, Sup 2 (for supplementary); S1, U1, S2, U2, S3, U3, U4, S4, U5, S5 (for the initial measures of ability with the skilled and unskilled hands), S6-S40 * (for the continued practice with the skilled hand), and U6, U7, U8, U9, U10 (for the final measurements with the unskilled hand). Divide the table into blocks by double lines below P2, Sup 2, S5, and S40.

Thumb tack the star pattern or hold it steady on the table with the starting point directly in front of S. Read S the following instructions and make sure that he understands them.

"*Start at the point indicated and trace in the pathway of the star until you return to the starting place. Keep exactly in the path, and try not to let your pencil point even touch the lines which bound it. If you do cross the boundary line, never continue forward but trace back at once to the point where the error occurred, and then continue forward so as to leave a continuous line in the pathway. Do not hesitate but keep your pencil moving all of the time. Make the trip through the pathway as rapidly as you can consistently with the avoidance and correction of errors.*"

Throughout the experiment have S trace *counter-clockwise* with the *right*, and *clockwise* with the *left* hand. Be sure that you *label on every star pattern* the trial number to which it corresponds.

Preliminary.—Take two trials *without mirror or screen.* Record the time in seconds and record notes and remarks in rows P1 and P2. (Leave the Sup 1 and Sup 2 rows to be filled in later.)

* Twenty-five instead of 35 practice trials may be taken here to shorten the experiment somewhat. If this is done, change text to correspond to shortened series.

Initial Measurements.—Adjust the shield above the pattern and adjust the mirror beyond it in such a way as to prevent S from seeing the pattern or his pencil point directly, but so that he can see both pattern and pencil point indirectly in the mirror. Make sure that the screen will not interfere with the movements of the pencil in drawing. Take five trials with the skilled, and five with the unskilled hand in the order indicated under "Trial No."

Record in appropriate columns of your table S's times and all of his volunteered remarks which throw any possible light on his attitudes, thoughts, plans or tricks, direction of attention, etc., as was done in the maze experiment, and add any notes of your own which may help as explanations of changes or of unusual records.

Practice.—Take trials S6-S40, the mirror and screen remaining adjusted as for the initial measures. Record times, notes, and remarks as usual.

Supplementary.—Immediately after S finishes trial S40, have him repeat the preliminary exercise (no mirror or screen). Find out from him whether or not he finds introspective evidence of a tendency for the mirror habit to persist. Record results in rows Sup 1 and Sup 2 which were previously left vacant.

Final Measurements.—Replace mirror and screen and take trials U6-U10, recording times and notes and remarks.

Get from S, by questions if necessary, a general introspective account of the learning process. Find out especially whether he can tell you about (1) any definite *means* whereby he helped himself in learning, and (2) any definite *changes* in method or any changes in consciousness or in degree of attention given to par-

ticular aspects of the situation or problem, etc., as learning progressed.

Errors are to be computed as follows: As long as the pencil line keeps to the space between the two parallel lines no error is committed. Whenever the pencil line crosses one of the parallel lines just enough to be perceptible on the *outside* (toward the periphery or toward the center of the figure) an error is committed. Each time this happens one error is recorded. No new error is made until the pencil tracing returns to the space within the parallel lines and then goes out again. E should not count errors as S draws, but should label the papers (i.e., U4 or S4, etc.), and lay them aside and count the errors after the experiment is completed.

Results and Discussion.

(1) Summarize S's remarks and your observations of S's behavior, and state what light they throw upon the questions (a) as to how seriously S tried to avoid making errors, (b) whether or not he worked up to the limit of his ability, (c) whether or not definite plans, tricks or schemes were adopted, and (d) any general tendency to change method or attitude as learning progressed.

(2) Lay off a base line lengthwise of the cross section paper. Plot learning curves to represent the times for trials U1-U5, S1-S40 and U6-U10, grouped as indicated. Make breaks in the abscissa scale and in the curve between each two groups, thus indicating that trials with different hands are represented. By reference to S's behavior and verbal report try to explain any great irregularities in the curve.

On the same base with the above curves plot in dotted lines for comparison, a curve to represent the

following data obtained from a class of 60 students by taking the median time for each trial.

	U					S				
Trial No. ..	1	2	3	4	5	1	2	3	4	5
Sec.	210	143	120	95	92	228	122	95	85	76

	S									
Trial No. ..	6	7	8	9	10	11	12	13	14	15
Sec.	71	68	63	62	60	55	56	55	53	50

	S									
Trial No. ..	16	17	18	19	20	21	22	23	24	25
Sec.	50	48	47	46	45	48	46	46	45	43

	S									
Trial No. ..	26	27	28	29	30	31	32	33	34	35
Sec.	42	42	41	42	40	39	38	37	37	37

	S					U				
Trial No. ..	36	37	38	39	40	6	7	8	9	10
Sec.	36	36	36	35	34	58	55	54	50	50

Compare this curve with your subject's curve. If that for your subject differs greatly, what factors do you think were primarily responsible for the difference—a true difference in the ability of your subject, a difference from the class in his set (for speed, for accuracy, his persistence, interest, etc.), or some other difference? Justify your judgment by refer-

ence to the curves and to the results summarized in (1) above.

(3) Students are variable in the number of errors shown, but errors are usually so few that graphic representation is best made on the basis of grouped trials. Get the medians of the numbers of errors made in trials U1-U5; S1-S5; S6-S10; S11-S15; S16-S20; S21-S25; S26-S30; S31-S35; S36-S40; U6-U10. Use these medians as ordinates and plot a bar diagram to represent them. Keep the bars 3 times their own width apart and let the first and last bars stand still farther from the middle series of bars.

Alongside each of the bars of the above diagram represent by cross hatched or blackened bars the corresponding data for errors from the class: U1-U5 median = 10; S1-S5 = 7; S6-S10 = 5; S11-S15 = 5; S16-S20 = 5; S21-S25 = 4; S26-S30 = 4; S31-S35 = 3; S36-S40 = 2; U6-U10 = 6. Compare individual with class results and explain.

(4) Can you say from inspection of the curves whether practice with the more skilful hand has resulted in increased time-efficiency with the unpracticed hand? To be more certain compare the *average time* in trials U1-U5 with the average time in trials U6-U10.

How does this difference compare with the difference between the average time of trials S1-S5 and the average time of trials S6-S10? What conclusions do these comparisons suggest?

(5) What factors help to explain the phenomena of transfer? How do they operate? (See Pyle, pp. 293-319, esp. 310-317; or Jordan, pp. 192-216, esp. 213-216; or Starch, pp. 219-246, esp. 242-246.)

(6) What application has the principle of transfer of training in the field of education? (See Starch, pp. 247-295; or Jordan, pp. 219-247.)

(7) Judging from S's introspections and from a comparison of the times for P1 and P2 with those for Sup 1 and Sup 2, is there evidence that at the end of practice the new habit has to some extent interfered (at least temporarily) with the ingrained normal habit? Explain.

(8) Criticize the method employed in this experiment and suggest ways of improving the experiment.

Summary and Conclusions.

References

For a popular account of the results of experiments involving inversion of the retinal image, see G. M. Stratton: *Experimental Psychology and Its Bearing Upon Culture,* 1914, esp. pp. 142-151.

For mirror drawing in particular, see G. S. Snoddy: *Psychological Review Monographs,* No. 124, 1920.

For transfer of training results and discussion, see A. M. Jordan: *Educational Psychology,* 1928, pp. 192-247; D. Starch: *Educational Psychology,* 1926 (Rev. Ed.), pp. 219-295; W. H. Pyle, *The Psychology of Learning,* 1928 (Rev. Ed.), pp. 293-319; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 353-355; R. Pintner: *Educational Psychology,* 1929, pp. 263-284.

CHAPTER XX

Experiment No. 16

METHODS OF INVESTIGATING MEMORY *

I

The first experimental study of memory and of its dependence upon the various conditions of learning was made by Hermann Ebbinghaus in 1879-84. In some experiments Ebbinghaus used as units stanzas of poetry, choosing such as were of as nearly equal difficulty as he could find. In order to control more strictly the conditions of learning, however, he also invented and used as material series of "nonsense syllables," that is to say, combinations of letters consisting typically of a vowel or diphthong between two consonants. All units of such a series possess the merit of being of approximately equal difficulty for the individual learner, as well as fair for the different learners. Their use makes it furthermore possible to stop short without disturbance at any point in a mass of material and still to make reliable comparisons between the results obtained, not only by different observers, but also by the same observer at different times. Further to increase the reliability, such experiments were repeated many times over and conclu-

* An experiment in the psychology of testimony such as one of those described by G. M. Whipple (*Manual of Mental and Physical Tests*, 1915, pp. 383-408) is suggested as a supplement to experiments Nos. 16 and 17. Relearnings of the maze and of mirror drawing after a month or more are also valuable supplementary experiments.

sions were drawn only from averages. Yet greater reliability is secured by presenting the syllables or other units to the subject in a machine, so that each syllable appears by itself and for a constant and brief time only. S is instructed not to think of the material in the interval between learning and recalling, and various other rules are added to suit the particular conditions.

In the investigation of memory, several somewhat different "methods" may be followed. They differ chiefly in regard to the mode of learning and in regard to the mode of measuring the goodness or strength of memory. The following summary (pp. 244-245) of a number of such methods will explain itself upon careful study.

In order to secure reliable results by any of the above methods: (1) S must be *practiced* until he obeys the complicated instructions easily, and until he settles down to a moderately constant method of learning. In learning nonsense syllables, for example, he must be practiced until his mode of learning, his attitude (degree of attention and interest), and his emotional attitude become relatively stable. (2) The experiments must be conducted over a *long space of time* so that averages may be secured from many repetitions of the same experiment. (3) *Fatigue* must be avoided by learning in short and in properly distributed intervals. In experiments on forgetting, e.g., days and even years may sometimes elapse between learning and recall.

Since strict investigation requires such rigorous control and extended trials, the following experiments can serve only to demonstrate a few principles of method, and to measure very roughly and inaccurately some of the numerous conditions and results of memorizing. In particular, they attempt the *demonstration of the*

methods of complete mastery, prompting, right associates and retained members, using *various sorts of material*, and comparing the memorial results using sense *vs.* nonsense, direct *vs.* indirect associations; learning by wholes *vs.* by parts; and learning by "distributed" *vs.* "concentrated" repetitions.

The purpose of this experiment is: (1) to contrast the learning of sense *vs.* nonsense material using the method of complete mastery; (2) to illustrate the method of prompting in learning a vocabulary, and (3) to note the influence of intention to remember on learning and the effect of primacy and recency on memory.

Apparatus.—Mimeographed sheets and cards of materials for memorization and a mimeographed copy of special instructions for vocabulary test (see *Notes for Instructors*). Timer.

Method.—If time permits, each S may be required in each part below to learn four sections; thus completing an ABBA order, to compensate in part for practice and fatigue and to increase the reliability of results. Otherwise half of the subjects may learn sections AB, and the other half sections BA for each part, and the results then combined for a class record.

Part 1. *Complete Mastery: Sense vs. Nonsense.*—Secure from the instructor a sheet of learning material. Fold so that the first (or third) section can be presented alone.

Instruct S: "*Learn this material by wholes. Study by repeating it in a low voice with a definite and pronounced rhythm, and at what you consider the best or most effective rate.*

"*After the first repetition or two try as much as you can to 'anticipate' or to recite without looking at the*

Method	Mode of Learning
Retained Members	Controlled presentation till series is incompletely learned.
Right Associates	The first unit of each succeeding pair is accented (trochaic rhythm).
Complete Mastery	Series repeated till S can reproduce it entirely.
Saving	Series learned to complete mastery; after intervals relearned.
Prompting	Controlled presentation till series is incompletely learned.
Memory Span	Series of gradually increasing length each presented once.
Reconstruction	Controlled presentation till series is incompletely learned.
Recognition	Controlled presentation till series is incompletely learned.

sheet, but whenever you hesitate even momentarily glance at once at the sheet for help, and proceed without delay.

"Continue reading the material over and over until you are able to recite it completely without once glancing at the sheet. Is it clear exactly what you are to do?"

Lay the sheet before S and start the timer. During learning, make pencil strokes in your notebook, thus recording the number of repetitions. Make sure that

Measure of Memory	Special Advantages
Number of *items* correctly reproduced.	Can be easily used with groups of subjects.
Percentage of correct responses, when only the first unit of each pair is again presented.	Studies "false" responses and measures the strength of individual associations. Responses may be timed.
Number of repetitions required (or amount of time expended).	Measures verbatim memory.
Number of repetitions (or time saved) in *relearning*.	Measures the strength of associations too weak to bring about unaided reproduction.
Number of prompts (helps) later necessary to bring about one or more complete reproductions.	Discovers the position of and measures the strength of the weakest associations.
Number of units in the longest series which can be *entirely reproduced*.	Is one of the simplest methods to apply.
Number of and "size" of *mistakes in reconstruction* of original seriès, when later its units are presented in haphazard order.	Can be used with non-verbal material (e.g., smells). Studies place associations.
Percentage of *correct recognitions* when original and similar units are later presented.	Makes possible a comparison of recognition with recall.

S obeys instructions, particularly with regard to learning by wholes and not repeating parts by themselves. At the end of complete learning, stop the timer and record the total time elapsed and the total number of repetitions taken.

After a minute of rest, have S proceed to learn the next section of the material in the same way at his own best rate. Keep similar records of number of repetitions and of time elapsed.

Part 2. *Prompting: Learning a Vocabulary.*—Return the material and secure special instructions and a sheet containing two Turkish-English vocabulary lists. Fold so that one list only is visible. Make out two tables of 20 rows each and with 2 wide and about 20 narrow columns. Copy the lists of Turkish and English words in the first 2 columns of the tables.

Instruct S: "*Learn this vocabulary by pronouncing each pair of words once slowly and in a low voice. Read at about this speed* (show by tapping on the table every two seconds for a word, or about four seconds per pair).

"*Many persons find it helpful to attempt to make artificial associations between the words.*

"*After one reading, I shall cover the English words and ask you to give them to me as you look at the Turkish words in order. When you reply incorrectly or hesitate longer than three seconds, I shall prompt you and pass on to the next word.*

"*After this we shall continue trials with prompts until you perform a complete trial with no prompts whatever.*"

Show S the first list and start the timer. In the first trial, be sure that S reads each pair aloud slowly and but *once*. Let him take no more than four seconds for any pair. Then cover the English words, and let S henceforth see only the Turkish.

In the succeeding trials, as soon as S gives the correct response (or as soon as his three seconds are up and you have prompted him) go on to the next word. Do not let him repeat your prompting word or your correction more than once.

Record 20 prompts for trial 1 (for reading through the list is equal to a prompt for each word), and for the

succeeding trials record in the appropriate column and row an "X" for each prompt and for each corrected error.

At the end of the final (perfect) trial stop the timer and record the total time elapsed. Ask S to tell you as best he can what helps or "tricks" he used in learning.

Three minutes after this test has been finished, give S a special test, the directions for which the instructor gave you when you got the vocabulary lists. Record in a new table the total time of test, and notes on S's replies and failures.

Instruct S: *"You are now to learn this second list under exactly the same conditions as you did the first one, except for practice and for knowledge that a special, as well as a regular test will be given. You will be given the same kind of special test as with the first list of words."*

Again record the prompts and the total time for the regular series and for the special test. Ask S to report what, if any, difference in method of learning occurred with the second list. Note any disturbances, volunteered remarks, etc., which may throw light upon results.

Results and Discussion.

Part 1.

(1) How many repetitions did S require for complete mastery of the verse *of poetry?* For the verse of *nonsense material?* How many seconds for each? Compare the learning of sense with nonsense material and try to give reasons for any differences found.

(2) Compute how many times harder it is completely to master a verse of the nonsense than one of the sense

material as measured in *repetitions,* as measured in *time required.*

Ebbinghaus, using this same method with long lists of one syllable nonsense material and long selections from this same poem (Byron's *Don Juan*) found nonsense took about *ten times* as many repetitions as sense. How do your results differ from those of Ebbinghaus? Give one or two important reasons why they differ from his.

(3) Give two important reasons why nonsense material is harder to learn than sense (for help, see Woodworth, pp. 77-78).

State one important drawback (or disadvantage) of the method of complete mastery.

Part 2.

(4) Compute and record the total number of prompts required in each successive trial in learning the Turkish-English vocabulary of 20 words the first time. Plot a learning curve from the data. Do you consider it a typical learning curve? Why?

(5) Compare the total time of learning the first with the total time for the second list. Also compare the number of trials before complete mastery was obtained in each of the two lists. Try to explain any marked differences found between them.

(6) Summarize the results of the special test with the first list; with the second. Score each in terms of number right plus half of number half right. (Half credit is given for responses in which *half* or *more than half* of the letters are correct.) Compare the results of the two mathematically. Which list gives the greater proportion of fully correct responses? Of half right responses? Try and explain any dif-

ferences found. State verbally what the results show as to direct and indirect association and the intention to remember (spelling and paired associates irrespective of position in series).

(7) For each list of words find the number of prompts for the first word in the series, the average number of prompts per word for words 2 to 19 (inclusive), and the number of prompts for the last word in series.

In relation to learning, the law of *primacy* says that, other things being equal, the first association is most likely to be recalled. The law of *recency* states that, other things being equal, the most recent association is most likely to be recalled. Relate your results in the first part of this question to these laws of association citing figures to support your interpretations.

(8) Criticise the controls used in this experiment and suggest changes in method to improve the certainty of results in such an experiment.

Summary and Conclusions.

References

See conclusion of experiment No. 17.

CHAPTER XXI

Experiment No. 17

METHODS OF INVESTIGATING MEMORY

II

The purpose of this experiment is: (1) to compare economy in distributed *vs.* concentrated learning of a vocabulary by the method of right associates; and (2) to compare the economy of whole *vs.* part procedure in learning poetry by the method of retained members.

Apparatus.—Cards on which are printed pairs of Chinese words and their English equivalents. Sheets containing selections of poetry. (See *Notes for Instructors.*) Timers.

Method.

Part 1. *Right Associates: Distributed vs. Concentrated Learning of a Vocabulary.*—The instructor will act as experimenter and the entire class as subjects.* Arrange the seating so that the printed cards can be easily read by all. Warn the subjects not to whisper; not to glance at another's paper, etc.

Instruct them: "*I shall show you, at the rate of one pair per three seconds, a list of twenty-four pairs of Chinese and English words. Read them silently, mentally accenting the first member of each pair. Study each pair as it appears and try to connect the Chinese with its English equivalent.*

* This can be used as an individual experiment for groups working in pairs if the stimulus material is printed on small cards or mimeographed.

"*A few minutes after this learning, the Chinese words will again be shown to you in a different order and you will be called upon to recall the corresponding English words.*"

Give one repetition—rest three minutes, then one more repetition—rest three minutes; then one more—rest three minutes. Then show the 24 Chinese words one at a time in a different and predetermined order, allowing six seconds for recall in each case. Students record responses in a table with numbers 1-24.

A second list is then learned in the same way as the first, except that the three repetitions are given one after the other with no intermissions. After the third repetition, wait three minutes and then test as before.

If time permits, a third list is learned as was the second, and a fourth as was the first. All replies are then checked for correctness, by reference to the original pairs. Scores are recorded in terms of number or percentage of correct responses.

Part 2. *Retained Members: Wholes vs. Parts in Learning Poetry.*—Obtain from the instructor either sixteen verses of "Alice Brand" (sectioned $2 + 2 + 2 + 2$ and 8) or sixteen verses of "John Gilpin's Ride" (sectioned 8 and $2 + 2 + 2 + 2$). The former is to be learned half by parts, then half by wholes; the latter is to be learned half by wholes, then half by parts.

In learning by WHOLES, tell S: "*Study the poem as a whole at any rate you like for twelve minutes. I shall warn you after six, and again after ten minutes. If a part seems hard you may take it more slowly, but don't repeat any single word or phrase or line or verse over and over by itself. Go straight through the whole selection from end to end each time, reading in a low voice,*

so that I can count the number of repetitions. You will do well to try to recite as you learn, but don't wait long before glancing back at the paper when you hesitate.

"After learning I shall give you a three-minute rest in which you are not to think about the poem. Then I shall call upon you to recite, and prompt you as many times as necessary."

Fold the sheet if necessary in order to conceal all except the section S is to learn. Give the sheet to S and start the timer. Count the repetitions by pencil strokes in your notebook. Give the proper warnings to S at six and ten minutes, and stop him at the end of twelve. After three more minutes, call upon S to recite. When he goes wrong correct him by three words. When he hesitates for as long as ten seconds prompt him by giving three words.

Record the total time required for recitation (including prompts). Record the errors by O, O, O, and all full prompts by X, X, X. Keep the prompt records for the different verses or sections separated and clearly labeled.

In learning by PARTS, tell S: *"Study this section, reading it over and over at any rate you think best. At the end of three minutes I shall stop you. Then we shall take the other three sections one after another in the same way and spend three minutes in learning each one. Read each over and over in a low voice so that I may count repetitions. You will do well to try to recite as you learn but don't wait long before glancing back at the paper when you hesitate.*

"Three minutes after you finish the last section, I shall ask you to recite, and shall prompt you whenever necessary."

Fold the sheet to conceal all except the section S is learning. Give it to S and start the timer. Count S's repetitions by pencil strokes in your notebook. Stop S at the end of three minutes, and at once give him the next section to learn. After finishing the final section let S rest three minutes, trying not to think of the poem. Then ask him to recite all four sections. Correct him when necessary by giving three words. When he hesitates for as long as ten seconds prompt him by three words.

Record the total time required for recitation (including prompts), and for each verse or section separately record the total errors and the full prompts by O, O, O, and X, X, X, respectively.

Results and Discussion.

Part 1.

(1) Compare the number of *correct replies* given three minutes after the vocabulary was learned by *distributed repetitions* with the number after it was learned by *concentrated repetitions*. Explain.

What factor favoring forgetting is greater in the case of material learned by distributed repetitions? This factor tends to cancel or to outweigh the advantage of distributed learning. Does it do so in your results?

(2) (To be answered only if all four vocabularies were learned.) State mathematically the apparent influence of practice in your results in learning vocabulary under the above conditions.

(3) What have experimental results shown concerning distributed *vs.* concentrated learning? (See Jordan, pp. 164-169; Perrin and Klein, pp. 271-274;

Sandiford, pp. 220-225; Pyle, pp. 47-53; Gates, pp. 359-363; or Starch, pp. 167-190.)

What practical application has this principle in education? (See the above Gates or Jordan reference.)

Part 2.

(4) Compare the number of repetitions given the materials learned as a whole with the number in part learning. State what the comparison tends to show.

(5) We shall count an error which the subject corrects himself (O) only one-half as serious as a prompt (X). In such terms express the total prompts given in the recitation of the materials learned by wholes (e.g., $32 + 6/2 = 35$). Make a similar computation for the materials learned by parts. Which procedure then, judged by your results, appears to be more *economical* (requires least prompts) under these conditions? Explain.

(6) Summarize previous results on whole *vs.* part learning and compare these results with your findings. Give two reasons why learning by wholes usually proves economical. What bearing do these results have on methods of study? (See Jordan, pp. 171-175; Sandiford, pp. 225-227; Perrin and Klein, pp. 274-280; Gates, pp. 363-367; or Woodworth, pp. 85-87.)

(7) Outline a "mixed method" for learning poetry which will utilize both whole and part learning to the best advantage, i.e., not learning by parts nor learning by wholes entirely but by a method intermediate between them. (See references in question 6.)

Summary and Conclusions.

References

R. Pintner: *Educational Psychology,* 1929, pp. 224-250; A. M. Jordan: *Educational Psychology,* 1928, pp. 164-175; P. Sandiford: *Educational Psychology,* 1928, pp. 220-235; R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 72-123; D. Starch: *Educational Psychology,* 1927 (Rev. Ed.), pp. 167-190; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 355-385; W. H. Pyle: *The Psychology of Learning,* 1928 (Rev. Ed.), pp. 40-53, 144-184; F. A. C. Perrin and D. B. Klein: *Psychology,* 1926, pp. 267-280; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 346-367; E. B. Titchener: *Text Book of Psychology,* 1911, pp. 380-389; H. Ebbinghaus: *Memory* (trans. by Ruger and Bussenius), 1913.

CHAPTER XXII

Experiment No. 18

RATIONAL LEARNING *

The ordinary learning experiment is concerned with practice effects which on the whole are uninfluenced by rational behavior. Card sorting, memorizing vocabularies, letter-digit substitution, memorizing nonsense syllables, and learning to typewrite are examples. As they are largely concerned with the acquisition of skill the processes require little organization or conscious selection.

Acquiring the mastery of puzzles and mazes represents another type of learning in which the situation reacted to sometimes changes as practice progresses. In mazes, for example, it is usually impossible to employ reason during the early trials because the subject cannot react to the numerous cues at the same time, nor can he keep them all in mind until several repetitions have been completed. However, a consistent method of attack is usually developed during the latter part of learning in a maze. In other problems, as in certain puzzles, one successful solution may be sufficient to render future trials practically perfect.

While some early workers such as Darwin and Romanes contended that animals reason, Lloyd Morgan

* The rational learning test is used with the permission of Professor Joseph Peterson. The discussion, contrasting rational learning with other kinds of learning, is taken largely from Professor Peterson's articles (see references).

answered them by stating that such learning is trial and error. Thorndike, interpreting the reactions of animals to various problem boxes, decided that the evidence supported the trial and error theory. More recently certain results, particularly those obtained by Gestalt psychologists, indicate the possible presence of ideation (insight) in the behavior of animals, especially higher animals such as chimpanzees.

A tendency in psychological writing is to put the emphasis on mere contiguous association in explaining learning. In these writings vividness and recency are employed to supplement frequency of contiguousness.

The rational learning test was devised by Joseph Peterson "to see how rational learning is related to learning that must depend wholly on 'trial and error' efforts. How effectively are ideas used in a type of learning in which their employment is obviously helpful?" The test requires the subject to associate in random order the numbers 1 to 10 with 10 letters of the alphabet. This problem is somewhat different from those which can be solved by usual associative learning. At first it is necessary to give a series of preferences which are necessarily guesses. The number of guesses may be greatly limited through a rational organization of the situation. Thus the subject in reacting to a changing status for each correct response changes the situation slightly by limiting the range of possible guesses.

The reliability of the Rational Learning test, when scores on one form are correlated with those on another, is .60 for repetitions, .70 for errors, and .75 for time scores. As an intelligence test the Rational Learning test has a validity coefficient of .37 (Stanford Binet

scores as a criterion). The correlation between Rational Learning and the Myers Mental Measure scores is .53.

The purpose of this experiment is to obtain an objective measurement of rational learning, to compare methods employed in rational learning with those in ordinary associative learning, and to determine and evaluate the reliability of the test.

Apparatus.—Timer, mimeographed sheets of fore-exercise, Test I and Test II of Rational Learning Test (see *Notes for Instructors*).

Method.—Throughout the experiment E should arrange his material so that S cannot see the mimeographed sheets and the scoring. The fore-exercise is given to acquaint S with the nature of the problem. (The instructor should demonstrate the method and scoring by giving a 5-letter form of the test to an assistant or one of the students before the class. Use a sequence of numbers that is different than in the fore-exercise. Write scores on the blackboard.) The subject is instructed as follows:

"*In this experiment the letters V, W, X, Y and Z are numbered 1, 2, 3, 4, and 5 but not in that order. They are numbered just by chance. I shall call out V and you are to guess until you find the number that goes with V. When you get it right I shall tell you and then call out W. After the first trial, instead of saying 'right' when the correct response is made, I shall merely call out the next letter in the series. When you have guessed the right number for W I shall call out X. In this manner Y will come next and then Z. After the right number for Z has been given I shall start with V again and continue in this manner until you can give two complete repetitions of the series correctly. All you have to do*

is to remember the number that goes with each letter so that you can give the complete series of numbers without making an error. A conscious effort to remember what number goes with each letter and clear understanding of the nature of your task will increase the speed of your learning. Remember that each letter has a different number. Your record is to be compared with those of the other members of the class."

When S is ready E notes down the exact time and calls out V, the first letter. If S stops after his first guess, tell him to go on until he gets the right number. The records will consist of time in minutes and seconds (from the calling out of the first letter until the last response is given), number of errors (wrong responses), and the number of repetitions until two correct ones are made in succession. Tabulate your records as in the following sample record: (Note that all responses, correct and incorrect, are recorded and that a horizontal line is drawn across the page at the end of each trial or repetition.)

	V	W	X	Y	Z	
Numbers assigned	2	1	4	5	3	
Responses in 1st trial . .	4	3	3	3	(6)	Tester says: "No, there
	5	5	4	5	3	are only 5 letters."
	2	1				
Responses in 2nd trial.	2	1	4	3	3	
				5		
Responses in 3rd trial.	2	1	4	5	3	
Responses in 4th trial.	2	1	4	5	3	

Summary: Time, 1′ 42″; errors, 8; repetitions, 4.

After the fore-exercise is finished and you have satisfied yourself that S understands the task clearly, obtain

Test I and Test II from the instructor and make out a blank form similar to the sample for recording responses. Then give Test I. Instruct S as follows: "*We shall now take all the letters from A to J (Test I) proceeding as in the fore-exercise. The numbers will go from 1 to 10.*" When Test I is completed give a rest of 5 minutes and then do Test II with exactly the same procedure.

The class data (time, errors, repetitions) for Test I and Test II are then tabulated on the blackboard to be copied by the students for their computations. All students are to write up this experiment.

Results and Discussion.

(1) For your S, is the time taken for learning Test II any shorter than that for Test I? Compare median times of the class on Test I and Test II. Explain. Using the time scores of the class compute the percentage of cases in Test I which reach or exceed the median in Test II. Interpret this result. How could you show this overlapping graphically?

(2) Compute the reliability of the Rational Learning test for the class data in the three following ways: correlate time scores of Test I with time scores on Test II; errors of I with errors of II; number of repetitions in I with those in II. Compare the relative value of the three different criteria of learning, i.e., time, errors, and trials. Which criterion of learning might be discarded? Why? Which is the best criterion and why?

(3) For a group of 49 college students the reliability coefficients for Rational Learning were: Time, .75; errors, .70; repetitions, .60. How do the reliability coefficients from your data compare with these typ-

ical results? If your coefficients are considerably lower, how might the difference be explained?

(4) How would you proceed to obtain a validity coefficient for the Rational Learning test as a measure of intelligence? As a measure of college aptitude? State clearly the distinction between validity and reliability.

(5) Plot the learning curve of S for errors. Try to explain any marked irregularities in the curve. Is this a typical learning curve? Why?

(6) Compare and contrast in detail the method of learning employed in this experiment with that used in ordinary associative learning as described by Woodworth, pp. 77-80, and as employed by you in "Memory Methods I."

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1921, pp. 333-338; A. I. Gates, *Elementary Psychology,* 1928 (Rev. Ed.), pp. 312-314, 346-348; Joseph Peterson: "Experiments in Rational Learning," *Psychol. Rev.,* 1918, Vol. 25, pp. 443-467; "Tentative Norms in the Rational Learning Test," *Jour. Appl. Psychol.,* 1920, Vol. 4, pp. 250-257, "Comparison of White and Negro Children in the Rational Learning Test," *The Twenty-seventh Yearbook of the National Society for the Study of Education,* 1928, Pt. I, pp. 333-341.

CHAPTER XXIII

CORRELATION, RELIABILITY AND VALIDITY

Meaning of Correlation.—The statistical measures employed up to this point have yielded knowledge of individual or group performance in some specific trait or ability. Comparisons have been made between measures of central tendency and between measures of variability. Frequently it is desirable to know what degree of correspondence exists between performance in one capacity and performance in others, or to what extent one may predict the scores or standing in one test on the basis of scores in another. For example, if a student stands high in arithmetic, does he also stand high in geography? If one has a low rank on an intelligence test does he rank low or high in history; or in general scholarship? In these cases and in many others like them it is highly important for us to know to what extent the scores or standing in one series of measures allow us to predict ranks in other kinds of performance or traits under consideration. The method of correlation is employed to determine the extent to which such predictions may be made. We can define correlation as a statistical device for ascertaining the extent to which scores or standing in one series of measurements (one variable) may be predicted upon the basis of known scores in another variable. The quantitative statement of this extent of predictability is called the coefficient of correlation and is designated, when computed by the rank method, by ρ (rho).

Suppose that we give an intelligence test to a large and "heterogeneous" group of persons, some having high, some average, and some low or very low intelligence. Suppose further that the relative degree of intelligence of these persons in the group is obtained from the combined estimates of teachers, friends, employers or others who have known them long and thoroughly. By these estimates, subject A is given rank No. 1 (most intelligent), B rank 2, C rank 3, etc.; Y the next to the bottom rank, and Z the lowest rank (least intelligent). If now, subject A gets the highest score in the intelligence test, B the next highest score, C the next and so on, with Y the next to the lowest score and Z the very lowest, we have a *perfect positive* correlation between test score and estimated rank, and if our group is a typical or unselected one, we can be sure that the test is good for testing what people call intelligence. Even if the correspondence between score and estimated rank is not perfect, if most persons maintain approximately the same relative rank the same conclusion can still be drawn, though with a lesser degree of assurance. However, if there is little correspondence between estimated rank and size of intelligence test score, we say that the correlation is low and the test clearly does not measure what the estimators consider intelligence to be. If there is no discoverable correspondence between test score and estimated rank there is a zero correlation (allows no prediction).

If subject A (rank 1) gets the lowest score, B (rank 2) the next lowest and so on, with Y obtaining the second highest and Z the very highest score, the test scores and the estimated ranks show a *perfect negative* or inverse correlation.

When the coefficient of correlation is computed by the Pearson product-moment method (r) or when calculated from ranks (ρ) the coefficient is +1.00 for perfect positive correlation, zero (0) to indicate no correspondence, and —1.00 for perfect negative correlation. In practice in psychology and biology perfect correlations are seldom encountered. There is usually some tendency for the measures of the two arrays to correspond, but rarely indeed is there either perfect agreement or perfect disagreement. The coefficient of correlation, therefore, usually appears as a positive or a negative decimal. Thus +.90 * indicates a marked tendency for the pairs of measures in two arrays to correspond in magnitude permitting quite accurate prediction of rank in one test from rank in the other; and +.10 indicates a slight tendency toward agreement, and —.75 a fairly strong tendency towards disagreement.

Computing Correlation.—One of the simplest formulæ for obtaining the coefficient of correlation is that used in the method of *rank differences:*

$$\rho = 1 - \frac{6 \Sigma D^2}{n\,(n^2 - 1)}$$

This formula takes account, not of the absolute differences in the magnitude of the measures compared, but only of the differences in rank between paired measures. In the formula, ρ (rho) represents the coefficient of correlation; Σ represents "the sum of"; D represents difference in rank; and n the number of pairs of measures.

To obtain the coefficient: (1) Arrange the arrays in two vertical columns with the two scores for each subject opposing each other. (For beginners confusion is

* The (+) sign is usually omitted.

avoided by placing the scores of one series in rank order, the *best score* at the top of the column. Then the ranks in this test will be 1, 2, 3, etc. The ranks in the second test will not follow this 1, 2, 3 . . . order except when there is a perfect positive correlation. Be careful to keep the two scores for any subject always opposite each other.) See illustrations following this discussion.

(2) In the third column write the rank for each measure of the first array, and in a fourth column the rank for each measure in the second array. (For very long arrays it is safest to use small slips of paper upon each of which a pair of figures is written; then to arrange the slips in rank order, first for one array and then for the other, entering individual ranks on each slip.)

(3) In a fifth column (D) enter the difference between the ranks for each pair.

(4) In a final column (D^2) enter the squares of each individual D.

(5) Add the figures of the D^2 column.

(6) Multiply this sum by 6.

(7) Find $n\,(n^2 - 1)$.

(8) Divide the product obtained in (6) by the product obtained in (7).

(9) Subtract the quotient obtained in (8) from 1.00 to obtain the coefficient desired.

To represent the correlation graphically, draw a square, n units on a side. Represent each pair of ranks by a point whose abscissa is the rank in one array and whose ordinate is the rank in the other. Either letters or large dots may be used for these points on the graph.

The meaning of correlation, the procedure and arrangement of computation, and a method of graphic

representation of correlation are illustrated by the solution of the following short problems:

PROBLEMS ILLUSTRATING THE COMPUTATION OF ρ

Person	Score in Test A	Score in Test B	Rank in Test A	Rank in Test B	D	D^2
A	117	60	1	1	0	0
B	102	50	2	4	2	4
C	98	49	3	5	2	4
D	94	55	4	3	1	1
E	83	58	5	2	3	9
F	80	41	6	7	1	1
G	77	46	7	6	1	1

$n = 7$ $n^2 = 49$ $(n^2 - 1) = 48$ $\Sigma D^2 = 20$

6

$n(n^2 - 1) = 7 \times 48 = 336$ $6 \Sigma D^2 = 120$

$$\rho = 1 - \frac{6 \Sigma D^2}{n(n^2 - 1)} = 1 - \frac{120}{336} = 1 - .36 = +.64$$

Person	Repetitions Taken to Learn	Score in Vocational Test	Rank in Ability to Learn	Rank in Vocational Test	D	D^2
A	11	93	1	1	0	0
B	13	60	2½	6	3½	12¼
C	13	71	2½	4	1½	2¼
D	14	60	4	6	2	4
E	16	76	5	3	2	4
F	17	80	6	2	4	16
G	20	60	7	6	1	1

$\Sigma D^2 = 39\frac{1}{2}$

6

$6 \Sigma D^2 = 237$

$$\rho = 1 - \frac{6 \Sigma D^2}{n(n^2 - 1)} = 1 - \frac{237}{336} = 1 - .71 = +.29$$

(1) Notice that the *best score* receives rank 1, and that *highest score is not always best score.* For example, in

measurement of learning ability by repetitions or time taken or errors made, the lowest score is best, and is given rank 1.

(2) Notice that when two or more persons have the same score, they must tie in rank. Thus if two persons have the same score and thus tie for 2nd and 3rd places, each must be given rank $2\frac{1}{2}$, and the next person must be given rank 4 (see the 2nd illustration above). In the same way three people having the same score are tied

PERFECT POSITIVE CORRELATION

Person	Rank	Rank	D	D^2
A	1	1	0	0
B	2	2	0	0
C	3	3	0	0
D	4	4	0	0
E	5	5	0	0
F	6	6	0	0
G	7	7	0	0

$6\ \Sigma\ D^2 = 0$

$1 - \frac{0}{336} = 1.00$

ZERO CORRELATION

Person	Rank	Rank	D	D^2
A	1	5	4	16
B	2	6	4	16
C	3	1	2	4
D	4	3	1	1
E	5	2	3	9
F	6	7	1	1
G	7	4	3	9

$\Sigma\ D^2 = 56$

$6\ \Sigma\ D^2 = 336$

$1 - \frac{336}{336} = 1 - 1 = 0$

HIGH NEGATIVE CORRELATION

Person	Rank	Rank	D	D^2
A	1	6	5	25
B	2	7	5	25
C	3	4	1	1
D	4	5	1	1
E	5	2	3	9
F	6	3	3	9
G	7	1	6	36

$$\Sigma D^2 = 106$$
$$6$$

$$1 - \frac{636}{336} = 1 - 1.89 = -.89 \qquad 6\,\Sigma D^2 = 636$$

PERFECT NEGATIVE CORRELATION

Person	Rank	Rank	D	D^2
A	1	7	6	36
B	2	6	4	16
C	3	5	2	4
D	4	4	0	0
E	5	3	2	4
F	6	2	4	16
G	7	1	6	36

$$\Sigma D^2 = 112$$
$$6$$

$$1 - \frac{672}{336} = 1 - 2 = -1.00 \qquad 6\,\Sigma D^2 = 672$$

and must be given the same rank. Thus if three subjects tie for 1st, 2nd and 3rd places, each must be given a rank of 2 (the median rank of those tied), and the next person is given rank 4.

Notice also in this correlation that the number representing the rank of the lowest person is always n (the

same as the number of cases compared) unless the last rank happens to be tied with the next to the last.

(3) Notice that even with such fairly high correlations as +.64 there may very likely be considerable difference in rank in some individual cases, or small differences in a large number of individual cases.

This coefficient of +.64 does not mean "64% of agreement," as students sometimes assert. It means greater tendency to agreement than a correlation of +.60 would indicate, and less agreement than +.70 would indicate. Also a coefficient of +.64 does not mean just twice the amount of agreement signified by +.32. The +.64 simply means a greater amount of correspondence.

An important aid for interpreting closeness of correspondence denoted by a coefficient of correlation is the coefficient of alienation (k) which measures the absence of relationship between the two variables. The absence of relationship increases rapidly with decrease of the coefficient of correlation.* For theoretical or scientific purposes any correlation, large or small, may be of significance provided it is reliably established ($\frac{\rho}{PE_\rho}$ is large

* The coefficient of alienation (k) measures the lack of correspondence between two variables just as r measures correspondence. (The Pearson coefficient of correlation (r) may be deduced from ρ by $r = 1.0233\,\rho$.) It is obtained by $k = \sqrt{1\text{-}r^2}$.

The following table shows the rapid increase of k with decrease of r:

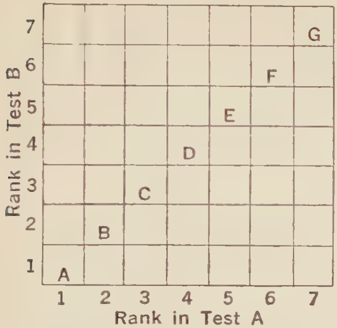
r	k	r	k
1.00	.00	.7071	.7071
.90	.45	.50	.87
.80	.60	.30	.95

Note that the amount of correlation which shows equal balance between the factors representing agreement and those representing disagreement is .707. Also note that the sum of these (.707 + .707) does not equal 1.00. See Kelley, p. 174 or Garrett, p. 290.

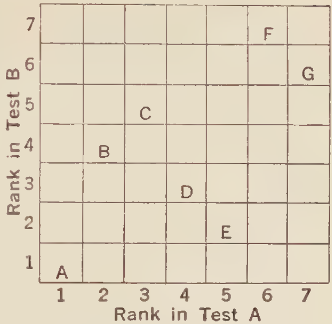
enough to assure high chances of confirmation on future experiments under like conditions). The size of correlation which is useful for predictive purposes in practical situations depends on the use to which it is put. Where measurements are rough and large units are employed, relatively low coefficients are meaningful. However if units of measurement are fine and great precision is demanded in determining ranks, coefficients should be relatively high for predictive purposes. There is no sharp division between a non-significant and a significant coefficient, i.e., the statement that .19 denotes a negligible and .21 a slight but significant correspondence is not true.

(4) Notice that a rough estimate of direction (positive or negative) as well as degree of correlation is possible from inspection of the graphic representation. With positive correlations representative points tend to cluster about a diagonal line across the square (from lower left to upper right corner), and the more nearly they approximate this diagonal the higher the correlation. With correlations near zero the points are irregularly scattered, while with negative correlations the points tend to cluster about the other diagonal. (Refer to above graphs.)

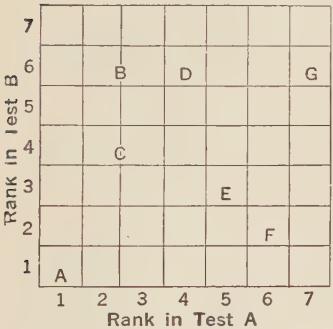
(5) The magnitude of a coefficient of correlation is usually influenced by the range of the measures dealt with, or, otherwise stated, by the degree of heterogeneity of the group. If, for example, the group is homogeneous, differences between scores will be slight. This increases the difficulty of accurate ranking, for a small change in score may put an individual into a different rank. Such difficulties, of course, tend to lower the correlation. With heterogeneous groups individual differences are larger, accurate ranking easier, and thus there is a tend-



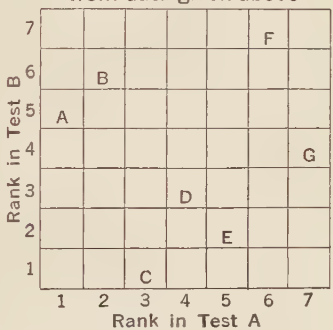
Graphic Representation of
Perfect Positive Correlation



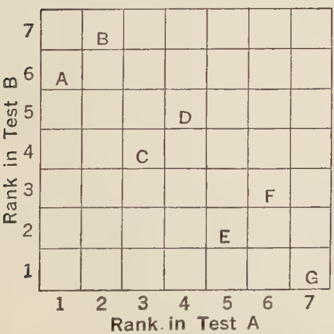
Graphic Representation of
Correlation of +.64
from data given above



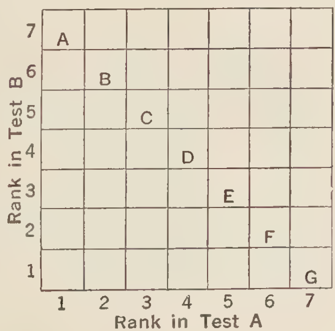
Graphic Representation of
Correlation of +.29
for data given above



Graphic Representation of
Correlation of Zero



Correlation of -.89



Correlation of -1.00

ency for coefficients to be higher. An example: For a group of college freshmen obtaining the highest 20 per cent of scores in an intelligence test the correlation between intelligence and school grades is .22 while for a random sample of the whole freshman class, in which there is a much greater range of scores, the correlation is .46.

(6) The assurance with which predictions or interpretations can be based upon a coefficient of correlation depends not only upon the magnitude and the reliability of the coefficient (explained in a later chapter), but also upon the degree of assurance that the group is an unselected one, and that the measures themselves are dependable measures of the capacities or qualities in question. Less dependable measures may be used for some purposes than for others, as we shall see later.

Application of the Correlation Method.*—There are a number of purposes for which the correlation technique is commonly employed.

(1) *Computing Reliability.*—Correlation is used to determine the consistency with which an instrument measures without regard to what it measures. This is frequently called the *reliability* of the measurement. Reliability, as the term is usually employed, is computed in different ways. Therefore it is important for the student to realize that what is predicted (by means of the reliability coefficient, see below) is dependent upon the method used for prediction, or in other words, that the so-called "reliability coefficients" do not, in all cases, mean the same thing.

* The instructor may omit discussion of reliability and validity to reduce the difficulty of statistical analysis. If this is done individual problems in experiment No. 18 and certain succeeding experiments must be omitted or modified.

Thus reliability of a measuring instrument may be considered as: (a) the extent to which scores or standing in a test which has been repeated (after a stated interval) may be predicted upon the basis of known scores obtained at the first testing (the *reliability coefficient* is obtained by correlating scores on test with those on retest); (b) the extent to which scores on a duplicate form of a test may be predicted upon the basis of scores on the first form; (c) the extent to which scores on odd items of a test may be predicted upon the basis of scores on even items; etc. Examples:

The Alpha intelligence examination was given to a group of 100 college students. One week later the very same test was again administered to the students. Correlation of scores on test and retest was .91. This means that the scores or standing on the Alpha test allow us to predict with a high degree of certainty the standing on a retest given one week later. That is, on the retest most of the subjects maintained relatively the same ranking that they had on the first examination.

Form 6 of the Alpha examination was given to a college group and one week later Form 8, which is similar and is considered to be equivalent to Form 6, was administered. Correlation between scores on Form 6 and Form 8 was .86. Thus, on the basis of scores in Form 6 one may predict with a fair degree of certainty the standing in Form 8 given one week later. Evidently the situation here is somewhat different from the above illustration. Corresponding items on the two forms of the test are probably not exactly equivalent.

A single form of Alpha was given to another group. The sum of the scores on the even-numbered items correlated .96 with the sum on the odd-numbered items.

Therefore scores on even items in Alpha are highly predictive of scores on odd items.

Considerable confusion arises from a loose use of this term "reliability." The reliability of a measure is frequently given without stating how the reliability coefficient was computed. Two reliability coefficients computed by different methods for the same test, or two computed by the same method for different tests do not mean the same thing.

The consistency of maze running or of card sorting for time scores may be obtained by correlating the sum of the times on odd trials with the sum of the times on even trials. The reliability of the error scores may be determined in the same way. These reliability coefficients merely show to what degree performance is consistent from trial to trial throughout the entire series.

The correlation coefficient which indicates satisfactory reliability depends somewhat on the nature of the test, and the size and variability of the group measured. Garrett considers that measuring instruments should have a minimum reliability coefficient of at least .80. Kelley states that a test should have a reliability of .90 or better if the results are to be used for individual comparisons; a reliability of .40 or better if group differences are to be investigated. Thus the minimum reliability permissible in a measuring instrument is determined partly by the use to which the obtained scores are put.

(2) *Computing Validity.*—Correlation is employed to determine the degree of correspondence between a test of a named trait and another test which is considered to be a more adequate measure of the trait in question. The *coefficient of validity* of a measuring instrument is computed by correlating its scores with those on another

measure called a criterion.* Thus a group intelligence test may be validated against: (a) scores on some other test considered to be a more adequate measure of intelligence; (b) teacher's ratings of intelligence; or (c) school marks. For example, (a) the validity of the Army Alpha when validated by means of the Stanford Binet test (as criterion) is .80 to .90, the size of the validity coefficient varying somewhat with the group used; (b) Alpha validated by means of teachers' ratings yields a coefficient of .67 to .82; and (c) with school marks the validity coefficient of Alpha is .50 to .60. These show that the validity coefficient obtained with one kind of "criterion" is different from that with another. Validity means no more than the dependability with which scores in one measure may be predicted upon the basis of scores in another measure which is correlated with the first.

Reliability then, as the term is commonly employed, means simply the correlation of a measure with itself or with another measure supposed to be identical with it. Validity means the correlation of a measure of a named trait or characteristic with another different measure which is supposed to be a more adequate measure of the same trait. It is obvious that "validity," as the term is commonly used, is a function of the adequacy of the criterion as well as of the measure to be validated. Essentially, therefore, validity means simply the correlation obtained by measuring or attempting to measure the same thing by two quite different methods and correlating the scores, while reliability is the correlation obtained

* The criterion is *assumed* to yield a true measure of the ability or performance in question. This rarely occurs in practice. In general, a criterion yields variable results on repetition of measurements and therefore cannot be a "true measure."

by measuring the same thing twice by the same or by highly similar methods and correlating the scores. In both cases the method or methods employed should *always be stated.*

(3) *Discovering Correspondence Between Traits or Performances.*—Correlation is employed to discover to what extent the scores in one trait or performance permit us to predict scores in another, for example: learning and quickness of forgetting; resemblance of children and parents in height, weight or intelligence; intelligence and memory; school achievement and intelligence; emotional stability and scholastic achievement; mental and physical traits of twins; speed and comprehension in reading; comprehension in reading and intelligence; etc.

Computation of Correlations.—Computation of correlations and probable errors of correlation coefficients will be facilitated by the use of Table 14 (squares and square roots), and Table 15 (to facilitate the computation of ρ) in the appendix of this book.

Exercises.—The following exercises are intended to give practice in the computation and interpretation of correlation coefficients.

(1) Find the correlation between chronological age and mental age for the following cases: Child A, chronological age, 4 years—mental age 5.7 years; B, 5—7.2; C, 10—12; D, 10—10.8; E, 7—6.8; F, 14—12.5; G, 10—9; H, 12—9.7; I, 12—10; J, 14—9.8; K, 12—8.6; L, 14—7.6. Represent the correlation of ranks in graphic form.

(2) Compute the correlation between I.Q. as determined by test, and the average of several teachers' ratings of intelligence for the following cases: Child A has an I.Q. of 1.40, and has an average rating by teachers'

estimates of 6.2; B, 1.40 and 8.2; C, 1.35 and 3.1; D, 1.20 and 2.6; E, 1.08 and 5.5; F, .97 and 8.6; G, .89 and 4.2; H, .83 and 5.5; I, .83 and 5.6; J, .70 and 8.2; K, .70 and 9.0; L, .54 and 11.3. Remember that an average teachers' rating of 1 is the highest possible rating, and that a high I.Q. indicates high intelligence. Represent the correlation of ranks in graphic form.

(3) Suppose that the following scores on an intelligence scale are obtained from persons of the ages listed—age 35, score 88; 40—88; 42—90; 48—84; 55—72; 58—76; 64—66; 68—66; 73—62; 78—60; 80—56; 86—54. What is the correlation between age and intelligence score for these older persons? What would such a correlation tend to show if the scale measures intelligence accurately?

(4) Suppose that in a cancellation test students whose average class standings were the following obtained the following ranks for accuracy:

Class Standing	*Rank in Test*
A	3
A	5
B	8
B	9½
B	14
C	9½
C	11½
C	3
C	11½
D	6
D	3
D	1
E	7
F	13

Judging from a correlation between class standing and rank in test, would giving the test to students at entrance to school be of use in predicting class grades to be expected later?

References

H. E. Garrett: *Statistics in Psychology and Education,* 1926; A. S. Otis: *Statistical Method in Educational Measurement,* 1925; T. L. Kelley: *Statistical Method,* 1924.

CHAPTER XXIV

Experiment No. 19

ANIMAL LEARNING *

All vertebrate animals profit by experience, that is, they learn. Their learning is fundamentally the same as that of man. They differ from man, however, in that their rate of learning is usually slower and the situations which they can learn are less complex. In order to secure learning in either man or animal there are certain conditions which must be set. In the first place the organism must be placed in a novel situation, that is, a situation to which the animal has not previously learned to respond. Secondly, the solution of the problem must lie within the ability of the subject (man or animal). Thirdly, the situation must be such that it will call forth a variable and analytical attack. In order to do this

* Investigation in the field of animal psychology was greatly stimulated by the early work of Thorndike. In 1896, working at Harvard, he experimented on instinctive and learned reactions of chicks. In this work he invented and used the maze for the first time. The next year Thorndike went to Columbia where he carried out his important studies of animal learning, using cats and dogs as subjects. The publication of these investigations in 1898 aroused much interest. Animal laboratories were almost immediately established at Clark and Harvard Universities, and a little later at the Universities of Michigan and Chicago. Since then studies in animal psychology have been steadily increasing. The white rat was first used in psychological experiments by W. S. Small at Clark University (published in 1899). Small was also the first to study the maze learning of these animals (published in 1900). Since these early studies the white rat has become one of the favorite experimental animals in learning investigations. This is because the rat is rather easily controlled in experimental situations and because it is a prolific and hardy animal.

there must be in the situation some sort of incentive or motivation. With animals the usual method of securing motivation is to control the food supply (producing hunger). When these conditions are met, we may observe during learning, the operation of the law of exercise and the law of effect in the weakening and elimination of certain connections, the selection and strengthening of others, and the combination of responses into complex series of reactions. The maze situation which fulfills the above conditions may be used for studying learning in rats.

Various investigations indicate that rats learn a maze first in terms of tactual and kinæsthetic cues. As the learning progresses, the control becomes more and more kinæsthetic and when the situation is thoroughly mastered the running may be considered as a kinæsthetic motor coördination with contact cues operating only infrequently. Watson has pointed out that rats deprived of vision, smell, hearing, or tactual sensitivity learn the maze about as readily as normal animals and hence kinæsthesis appears to function as the most important guide in learning this sort of problem. Other investigations show, however, that under certain conditions contact, olfactory, and visual receptors do play an important part in maze learning. Therefore, while such a problem can be solved by kinæsthetic factors alone, cues from other senses are probably used to supplement kinæsthesis in the learning of normal animals.

The learning of a simple maze by white rats will be observed in this experiment. Our purpose is to note the rats' behavior in mastering the maze situation and to make a quantitative determination of the progress made during the learning.

Apparatus.—A vertical maze (describe it), three white rats about 90 days old, a cage for the rats, two desk lamps to illuminate the front of the maze, food, a timer or stop watch. (See Appendix for diagram of maze.)

Method.—Before the experiment proper is begun the animals should be accustomed, so far as possible, to handling by the experimenter and to the room in which the experiment is to be conducted. This may be accomplished in the following manner: If the animals are entirely unaccustomed to handling, begin, about a week or ten days previous to starting the experiment, a daily routine of removing them from the cage and handling them for several minutes. If they are somewhat accustomed to being handled, a few days will suffice. For two or three days before the actual experiment is to be started, begin taking the animals daily into the room where the experiment is to be performed and where the maze is in the position which it will occupy during the experiment. Place each animal (separately) in the food compartment of the maze and allow several minutes for the rat to become acquainted with the compartment before introducing food and then feed the animal there. About eight minutes should be allowed for feeding before the animal is returned to its cage. In all probability the animal will not have spent the entire eight minutes in eating but will have eaten enough to maintain it until the next feeding.* After the first day, five or six minutes of feeding per day should suffice to keep the animal in a healthy condition and yet sufficiently hungry for food to act as an adequate incentive to produce learning. Careful handling and feeding of the ani-

* The animals' customary food will be satisfactory for use in the maze, i.e., bread and milk, grain mash, etc.

mals is *very important* if success is to be achieved in this experiment. While it is desirable that the experimenter be accustomed to handling rats, an individual without previous experience with animals can, if instructions are carefully followed, obtain satisfactory results.

Some experimenters follow the practice of forcing the animal to fast for a period of from 36 to 48 hours before the actual experiment is begun so that the hunger urge will be strong during the first trial. This practice accomplishes the purpose of securing strong motivation during the first trial of the experiment but at the same time is likely to upset the metabolism of the animal. Such an upset is undesirable in this experiment. If the procedure outlined above is followed and the food supply is cut down gradually, the metabolic disturbance is less severe and yet the motivation is sufficient to achieve the desired results on the first and later trials.

When the animal shows by its behavior that it has become accustomed to the food compartment, that is, when it goes to the food box immediately and starts eating when placed in the food compartment, the actual experiment may be started (using two of the three rats). Since the first trial is apt to be rather time-consuming it is advisable to run it while the class is working on some other experiment, or on the day before a regular class session if the class meets every other day. The data for this trial are given to the students at the next class meeting. On each succeeding day after the first trial the rats are given two trials per day for about seven days so that there will be a total of about 15 trials. On those days when the class does not meet the data on the maze running are collected by the experimenter and presented to the class at its next session. At least two rats should

be used for each laboratory section. On the day when these two rats receive their second and third trials allow the third rat, which has had the same preliminary handling and feeding but no maze running, to spend about 10 minutes in the maze to illustrate the exploratory behavior of a rat during its first trial. Remove this rat from the maze at the end of 10 minutes if it has not reached the food compartment. This animal is not used again in the experiment.

In this experiment the degree of learning is measured by the number of errors per trial and also by the time taken to complete each trial. (It is not likely that the maze will have been completely learned at the end of the experiment but the data collected should be sufficient to show very well the form of the learning curve. For our purposes the maze is said to be learned when 3 successive trials are completed without errors.)

The maze is so placed on a table at the front of the room that the open side is in full view of the class and with sufficient space in front of the apparatus for two desk lamps to be used for illuminating its interior. These lamps should be placed a short distance in front of the maze and far enough to each side to illuminate the maze satisfactorily without obstructing the view of the students. If relatively strong bulbs are used (100-watt), they will serve the dual purpose of making the reactions of the animal visible to those in the rear of the room and of cutting off from the animal visual stimuli from objects beyond the lamps. The nearest students should be at least ten feet from the front of the maze and the class should be cautioned to sit very quietly and not to make any sudden movements. All such stimuli will tend to excite and disturb the animal.

Before the maze running in class begins the instructor points out to the students the true path and the blind alleys of the maze. Previous to the first trial which the class observes, individual data sheets for each rat are prepared by the students and the records for the first trial (run the day before) are recorded on them. Each data sheet has four columns. Column one is headed "Trial No."; column two, "Time"; column three, "Errors"; and column four, "Comments." The individual members of the class are to count and record the number of errors made in each trial, and make notes on the behavior of the animal during each trial. They should pay particular attention to the changing behavior of the rat as the learning progresses from trial to trial. An error is committed by the rat every time the animal progresses far enough into a blind alley for his head and ears to show beyond the white lines painted on the wire front of the maze. The time is kept by the instructor and given to the students for recording at the end of each trial. An assistant counts errors as a check on the class.

After notifying the class which of the two rats is about to run the experimenter places food in the food compartment of the maze, carefully takes the rat from its cage, and places it in the starting compartment. As soon as the animal's body is in the maze, the instructor starts his stop watch and then stops it when the rat has completely entered the food compartment. As the rat goes through each trap door the experimenter releases the wire which allows it to close. This prevents retracing and consequently shortens the time required for the first few trials. At the end of the first trial for any day allow the rat to eat about 15 seconds and then proceed

with the next trial. When this is completed permit the animal to eat (in the food compartment of the maze) for about 5 minutes. No other food is given to the rats until their next trials.

After the first class session the two trials per day will consume only a few minutes at the beginning of each period. The rest of the session can be devoted to other experimental work.

Results and Discussion.

(1) Contrast in detail the behavior of a rat in the maze during the first few trials with his behavior during the final few trials. Explain the change. Is a rat's behavior in a maze ever entirely aimless or determined by chance? Why?

(2) Cite all the controls used in this experiment and explain why they are desirable. Criticise the method and suggest better controls.

(3) Construct a learning curve for one of the rats which finished the experiment using all his successive time scores as ordinates. Describe the curve and try to account for any distinct irregularities by relating them to the behavior of the rat as described in your notes. Is the learning curve typical? Why?

(4) Construct a learning curve on the same base line as in question (3) using average data from all rats run. (Use data from other sections if there are other laboratory classes.)

(5) Similarly draw a learning curve for one rat using the number of errors on successive trials as ordinates. Describe the curve and answer the questions asked in problem (3).

Construct a learning curve for errors on same base line using average data from all rats run.

If desired, the following results (group data from 6 rats for 15 trials) may be employed to plot curves from group data in (4) and (5):

Trial No.	1	2	3	4	5	6	7	8
Average time (sec.)	317	89	90	75	77	55	76	35
Average errors	16	7	6	6	5	6	5	4

Trial No.	9	10	11	12	13	14	15
Average time (sec.)	45	27	21	27	20	24	21
Average errors	3	3	2	3	2	3	2

(6) Compare the four learning curves in every way possible. Do you consider time or errors the better measure of learning in this experiment? Why? From examination of these curves state the effect on the form of the curve when the number of cases is increased. Describe a method which might be employed to smooth these learning curves.

(7) Cite and explain four prominent features which characterize animal learning in the maze (see Gates, pp. 308-310 for help in this question).

(8) Contrast and compare human learning with animal learning (for help in this, see Gates pp. 316-320; Woodworth, pp. 131-141; or Dashiell, pp. 332-333).

Summary and Conclusions.

References

R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 131-151; F. A. C. Perrin and D. B. Klein: *Psychology*, 1926, pp. 204-218; A. I. Gates: *Elementary Psychology*,

1928 (Rev. Ed.), pp. 307-320; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 325-333; J. B. Watson: *Behavior,* 1914, pp. 99-104, 198-219; M. F. Washburn: *The Animal Mind,* 1926 (3rd Ed.), pp. 246-337; J. B. Watson: *Kinæsthetic and Organic Sensations: Their Rôle in the Reactions of the White Rat to the Maze,* Psychol. Monog., No. 33, 1907; H. Carr: "Maze Studies with the White Rat." 1. *Normal Animals, Jour. Animal Behavior,* 1917, *7*, pp. 259-275; S. B. Vincent: "The White Rat and the Maze Problem," *Jour. Animal Behavior, 5,* pp. 1-21, 140-157, 175-184, 367-374; *The Function of the Vibrissæ in the Behavior of the White Rat,* Behavior Monog. Series, No. 5, 1912.

CHAPTER XXV

Experiment No. 20

LEARNING AND HABIT INTERFERENCE IN CARD SORTING *

When one habit has a negative effect on another, interference arises. The fumbling and unsuccessful moves which occur when a person accustomed to tying his bow-tie in front of a mirror attempts to tie the bow-tie of another standing facing him may be cited as an example of habit interference. This phenomenon of interference has been investigated in the psychological laboratory by sorting cards with the sorting boxes arranged successively in two different orders, by typing with two different keyboards, etc.

W. H. Pyle found that, in sorting cards according to two different layouts, it was more efficient to practice with one arrangement for 15 trials and then work with the other for 15 trials rather than to alternate from trial to trial. He concluded that if two complicated sets of mutually interfering habits are to be formed, the most economical procedure is first to form one set and then the other. Similarly J. F. Dashiell discovered that it was more advantageous to practice one habit until it was

* Card sorting was first used in psychology by J. A. Bergstrom in 1893 for the purpose of studying interference in habit formation. Joseph Jastrow, in 1897, recognizing in this type of test a technique of wide applicability, varied the method and enlarged the usefulness of the experiment. The possibilities of the test having become evident, card sorting, with many changes in details, came into extensive use in many later studies in the psychology of learning.

almost completely learned before attempting the other rather than to practice them alternately. This held for card sorting as well as for several other types of problems.

Two mutually interfering habits may exist side by side. While interference seems to be greater with fast than with slow learners, the former overcome such interference quickly. "It is a general finding that the interferences that seem striking when one or both habits are in their initial stages of formation, tend to disappear as the habits become well integrated." *

Ordinarily card sorting is done in the following manner: A pack of cards is held in the left hand, face down; the top card is removed with the right hand, turned over, and thrown into the sorting box corresponding to the suit (number, diagram, etc.), discovered on the card. The method to be used in the present experiment is somewhat different and may be called the "one-behind" procedure.† Holding the deck face down the top card is thrown, face up, in the upper left-hand box of the sorting compartments. The second card is thrown into the box corresponding to the suit showing on the face of the first card. In like manner the suit of card number two designates where the third card is to go, etc. Obvious advantages of this method over ordinary sorting are: (1) possibility of direct comparison of dealing and sorting; (2) better coördinated sorting movements; (3) less visual strain; (4) the more continuous nature of the task.

* See first Dashiell reference, p. 354.

† Used at Stanford University by G. M. Trace and later by R. W. Husband and W. R. Miles, "On Sorting Packs of Sixty Cards with Form and Color as Variables in Two to Six Kinds." *Jour. Appl. Psychol.*, 1927, Vol. II, pp. 465-482.

The purpose of this experiment is to study interference in habit formation by investigating progress in learning for card sorting; and to note the interference which occurs when the relative positions of the sorting compartments are changed.

Apparatus.—Timer or stop-watch. One pack of 60 playing cards with an equal number of cards in each suit (Jacks, Queens, Kings and Jokers removed). One sorting placard (see Appendix).

Method.—Place the placard with one margin even with the front edge of the table. Have S stand directly in front of the placard. Shuffle the cards, *see that the last card to be dealt is of the same suit as the upper left-hand compartment of the placard* (so that this compartment will contain only 15 cards when a trial is completed). Then instruct S: "*The sorting is to be done according to suits. Pay no attention to numbers on the cards. Note the locations of indices on the dealing placard. Hold the shuffled pack face down in the left (or unskilled) hand and when I say 'go' deal the top card face up to the upper left square. After the suit of the first card dealt has been identified, locate the index corresponding to the suit of this card, e.g., it may be the lower left-hand square. Then deal the second card face up to the square having the index just located (the lower left). When the suit of this last card is likewise noted, its corresponding index located, deal card number three to that square face up. In this manner continue sorting until the pack is exhausted. Should the card dealt have the same suit as the index of the square in which it lies, deal the next card to that same square, continuing so long as the same suit appears successively. The object is to achieve maximum speed with*

a minimum number of errors. Correct all errors that you notice." (The instructor should demonstrate the sorting to the class as a whole.)

Time S from the "go" signal until the fall of the last card, recording to fractions of a second if possible. Also watch for and record the number of uncorrected errors. If there are no misdeals there will be 15 cards in each square at the end of the trial. Complete 20 trials with about 45 seconds rest between successive trials (just time enough to reshuffle the cards and prepare for the next trial).

After the 20th trial has been completed give the sorting placard a ¼ turn to the right (clockwise). Now do 5 more trials of sorting and record times and errors as before. Take detailed notes on the behavior of S on the 21st trial, and contrast his behavior on the 21st trial with that on the 20th trial. Write down any voluntary remarks of S made during the 21st trial. Did S experience any "blocking" or inhibitions on the 21st trial?

Then list the class results on the blackboard. Record the time (only) for each trial, sum of the scores on the 10 odd numbered trials (scores for trials $1 + 3 + 5 + \ldots 19$) and the sum of the scores on the 10 even numbered trials for the first 20 trials. These data are to be copied by the students for their computations.

Results and Discussion.

(1) Compute the reliability of the sorting (first 20 trials) test as follows: correlate the total time for the 10 odd numbered trials with the total time for the ten even numbered trials (class data). How consistent is performance in card sorting from trial to trial throughout the series? Compare the coefficient

of correlation with its PE and state whether ρ is dependable.

(2) Plot learning curves for S to represent the times for trials 1 to 20 and 21 to 25, grouped as indicated. Make BREAKS in the abscissa scale and in the curve between the 20th and 21st trials to indicate the changed conditions under which the 2 groups of trials were given. On the same base line plot for comparison a dotted curve to represent the class data. (Use the median time for each trial.) Compare this curve with that of your S (for first 20 trials) and attempt to explain any marked differences present. State why the curves are or are not typical learning curves.

(3) By referring to the learning curves describe carefully the change in achievement in the 21st trial in comparison with the 20th. Account for this change. Summarize your description of S's behavior on the 20th and 21st trials and his remarks concerning inhibitions present. Relate this to the change in achievement from 20th to 21st trial.

Does the loss in skill revealed in trial 21 tend to be regained rapidly? Answer by noting the progress from trial 21 to trial 25 in comparison with the level attained at trial 20.

(4) Tabulate the error scores of S. On a similar experiment Husband and Miles state that errors are rare and inconsistent as to location in this card sorting test. Criticise this statement in the light of your results.

Compare by inspection the error and time scores of S, and state whether you consider error scores or

time scores the better measure of learning in this experiment. Why?

(5) Outline an experiment designed to compare learning in dealing cards (distributing into 4 piles in rotation), in ordinary sorting (looking at a card before throwing it into the proper compartment), and in sorting by the method used in this experiment. Cite controls and methods of handling data.

Summary and Conclusions.

References

W. H. Pyle: *The Psychology of Learning,* 1928 (Rev. Ed.), pp. 319-322; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 354-355; "A Comparison of Complete vs. Alternate Methods of Learning Two Habits," *Psychol. Rev.,* 1920, Vol. 27, pp. 112-135; W. H. Pyle: "Transfer and Interference in Card Sorting," *Jour. Ed. Psychol.,* 1919, Vol. 10, pp. 107-110.

CHAPTER XXVI

Experiment No. 21

THE SPAN OF VISUAL APPREHENSION

By the span or range of visual apprehension * is meant the number of disparate objects that can be recognized and named after visual examination of a series of items during a short period. The term "range of attention" has frequently been wrongly employed for visual apprehension. It is commonly stated in psychological texts that the range of attention for simultaneously presented objects is 4 to 6 items. Recent investigators have pointed out convincingly that such statements actually concern visual apprehension, not attention.

Much of the experimentation on range of visual apprehension has employed the tachistoscope, an apparatus for exposing stimulus material for relatively short intervals.†

Various factors influence the span of visual apprehension. The span increases with age, correlates +.60 to +.70 with intelligence, and span for words correlates +.70 with reading ability (speed plus comprehension). The type of stimulus material employed affects perception span. From 4 to 5 unrelated objects, letters, digits

* Also called perception span.

† Although the tachistoscope was used as early as 1859, the first important work on perception span in relation to reading appeared about 1885 when Cattell began publishing the results of his investigations. From then until about 1910 occurred a large amount of work along this line; since 1910, a smaller amount.

or words are apprehended per exposure. When objects are grouped in symmetrical patterns, symbols in formulæ, letters in nonsense syllables or words, and words in phrases or sentences the range of apprehension is increased. Thus context exerts a definite influence on span. There is a marked tendency to combine the different elements of a visual impression into higher perceptual units whenever grouping is possible.

Span is somewhat dependent on duration of exposure. With an exposure of $\frac{1}{400}$ second about 3.5 items of unrelated material are apprehended; with $\frac{1}{10}$ second, about 4.5 items; and with 6 seconds about 6 items. Practice increases the span markedly in the early periods of practice and slightly during the later stages.

The purpose of this experiment is to determine the spans of visual apprehension for letters in nonsense arrangement, for words in nonsense series, and for words in sentences. These results will be compared to demonstrate the effect of variation in the type of stimulus material used on the span of visual apprehension.

Apparatus.—Timer or stop watch; three lists of material printed on cardboard (see *Notes for Instructors*): (1) 10 series of letters in nonsense arrangement containing 10 letters each; (2) 10 series of words in nonsense arrangement containing 10 words each; (3) 10 series of words arranged in sentences of 10 words each. An exposure frame * for presenting stimuli (see Appendix) and a mimeographed key for scoring responses.

Method.—The instructor acts as experimenter and the class as subjects throughout the experiment. Distribute

* If it is not convenient to use an exposure frame the instructor can expose the stimuli by holding up the cardboards for the prescribed time. Timing of exposure is more difficult with this technique but apparently the results are not altered.

paper to the class and instruct the students to make an outline form for responses in the following order, writing the responses horizontally from left to right on the appropriate lines: 5 series of letters, 5 series of words, 10 sentences, 5 series of words, 5 series of letters. (The instructor is to explain the need for this ABCCBA arrangement.) After informing the subjects that 10 series each of letters, of words in nonsense arrangement, and of words in sentences are to be given, read the directions: "*The material will be exposed in the order given in your outline. Two seconds after the ready signal is given a card will be uncovered and remain exposed for 2 seconds. Look at the exposed material, reading it from left to right, and try to remember the items just as they are printed on the card. When I cover each series, but not before, you are to write down in the appropriate place on your paper the material just as it was given. Place a dash in the series to designate the omission of a letter or word.*"

Give the trial series, and have each student write his response on the back of his paper. Then, after ascertaining that all understand the task, give the 30 experimental series in the ABCCBA order described above.

Pass out the scoring key and have the students score their own papers as follows: Give a credit of 1 for every correct letter or word in the right position in the series, and a credit of ½ for each transposed item. Failure to indicate the omission of a letter or word transposes all succeeding items even though they are otherwise correct. Give ½ credit to each of these transposed items. The score for each kind of material is the average of the 10 series of that type. Place the average score of each S

on the blackboard for the class to copy. Each member of the class is to write up this experiment.

Results and Discussion.

(1) For the class data what are the standard deviations for each kind of material? The coefficients of variability? Which material shows the greatest absolute variability? Greatest relative variability? Why is it better to use the relative variability rather than absolute variability for such comparisons?

(2) Compare the average for isolated letters with that for words in nonsense arrangement. What are the chances in 100 that the discovered difference will be in the same direction in future experiments under like conditions? What does this signify concerning the reliability of the discovered difference? If for the same comparison you found $D = .20$ and $\sigma_D = .41$, how would you interpret your results?

(3) Similarly compare the average for isolated letters with that for words in sentences, the average for isolated words with that for words in sentences, and interpret your findings.

(4) Is it possible to obtain a difference between two averages that is perfectly reliable (exactly 100 chances in 100 that the difference will be in the same direction in future experiments under like conditions)? Why?

(5) In which cases are the differences less reliable: between letters and isolated words, between letters and words in sentences, or between isolated words and words in sentences? Why? Explain in detail.

(6) Give controls used in this experiment, stating why

they were employed and how adequate you consider them. Suggest, if you can, changes in method to improve controls.

Summary and Conclusions.

References

E. B. Huey: *The Psychology and Pedagogy of Reading,* 1907, pp. 51-101; G. M. Whipple, *Manual of Mental and Physical Tests,* Part I, 1914, pp. 262-296, Part II, 1915, pp. 150-204; M. A. Tinker, "Visual Apprehension and Perception in Reading," *Psychol. Bull.*, 1929, Vol. 26, pp. 223-240; F. N. Freeman: *Experimental Education,* 1916, pp. 110-116.

CHAPTER XXVII

RELIABILITY AND THE COMPARISON OF REPRESENTATIVE VALUES *

We have noted that owing to our incomplete control over the conditions of experimentation a single measurement is never sufficient, but that repeated measurements all made under a set of conditions held as rigidly constant as possible are necessary if we wish to obtain a *reliable* result. The question therefore arises: *How many times* must we repeat in order to secure *sufficient reliability?* Or *how reliable* is the result obtained from a given number of measurements?

The Reliability of Measures of Central Tendency.— It is clear from our discussion of variability that if our measurements all cluster closely about some single value, the chances are good that another similar series of measurements will give us an identical or nearly identical central value; whereas the greater the variability of the individual measurements the less the reliability or dependability of the central value. The reliability of an obtained average (or other measure of central tendency), therefore, bears an *inverse relationship to the measure of variability* of the distribution.

But our reliability or assurance depends also upon the *number* of measurements we take. Theory and practice

* If a more thorough treatment of this subject is desired the instructor should supplement the discussion in the text with lectures and illustrations. The instructor may omit Chapters XXVI and XXVII to reduce the difficulty of statistical analysis in the course. If this is done certain problems in a few of the succeeding experiments must be omitted or modified.

both show that the stability or reliability of an average is increased not exactly in proportion to the increase in number of measurements upon which it is based, but in proportion to the square root of that number. The average of 100 measurements is not 100 times, but only $\sqrt{100} = 10$ times as reliable as a single measurement. The average of 64 measurements is not 4 times as reliable as the average of 16 but only twice as reliable for $\sqrt{64} = 8$; $\sqrt{16} = 4$; and $8/4 = 2$.

The measures which we shall use to indicate the reliability of our measures of central tendency, therefore, will involve two factors; *the variability of the distribution, and the number of cases it contains.*

Now theoretically we never know and never can know absolutely the true value of anything, since all actual measurements are subject to error. The addition of even one new measurement to a given distribution will probably change the average a little, since it is quite unlikely that the single new measure will coincide exactly with the old average. The best we can do is to assume, where possible, that our particular set of measurements is typical, and not an "abnormal" or "specially selected" set of values, and then to inquire how greatly the averages of future measurements involving the same conditions and the same number of cases will probably differ from it. (Strictly speaking, how greatly an infinite series of such averages would vary from their own average. The latter is, of course, the theoretical "true" value.) This is the same thing, scientifically speaking, as to ask how nearly we have arrived at a "true value." If successive series of measurements of the same phenomenon were to be taken and successive averages obtained, it is clear that these averages would differ some-

what among themselves. If the original average was very reliable, they would tend to differ little; if the original average was very unreliable, they would tend to differ greatly. The original average therefore may itself be said to have a certain variability or tendency to vary, and a measure of the variability of the average is at the same time a measure of its *unreliability.* By making the assumptions that the original series is a typical one, and that the mathematical theory of probability can be applied for "predicting" future averages, we are able to derive the formulæ for unreliability given below. These measures of the variability of the average (σ_{av}, PE_{av}, PE_{med}) must be kept distinct from the measures of variability before considered, which are variations of the individual measures of the distribution *from* the average (or other central value). The measures of variability before considered are termed standard deviation of the *distribution* and probable error of the *distribution.* We shall designate them σ_{dis} and PE_{dis}, instead of plain σ and PE wherever a possibility of confusion arises. Measures of the variability of the average are termed *standard error of the average,* and *probable error of the average* (of the median, *probable error of the median*). They are designated by σ_{av}, PE_{av}, and PE_{med}. (Note that standard error and not standard deviation is employed for σ_{av}.) Hold the measures distinct by remembering that whereas the measures of variability *from* the average with which we are already familiar are obtained from *actual* measurements, these new measures of variability *of* the average are *theoretical* values based upon assumptions and referring to measurements *still unmade.* The former describe or represent actual conditions discovered; the latter *assume and predict* with a certain

degree of probability what is to be expected in the future. The former state a fact; the interpretation of the latter is conditional upon the validity of the assumptions underlying it.

The most commonly used measures of unreliability (variability of measures of central tendency) are the standard error of the average (σ_{av}); the probable error of the average (PE_{av}); and the probable error of the median (PE_{med}). Notice that their formulæ take account of the two factors already mentioned: the variability of the distribution and the number of cases.

$$\sigma_{av} = \frac{\sigma_{dis}}{\sqrt{n}}$$

$$PE_{av} = \frac{PE_{dis}}{\sqrt{n}} = .6745\sigma_{av}$$

$$PE_{med} = \frac{5PE_{dis}}{4\sqrt{n}} = \frac{5Q}{4\sqrt{n}}$$

When we are justified in assuming that our set of measurements is similar to that for the general population we can without further experimentation, apply the appropriate formula and state how reliable is our central value, or in other words how near our average or median probably is to the "true" one It is true also that unknown constant errors or unknown "selective factors" *may* have been operative in our original measurements. No amount of assumption and mathematical treatment can make up for lack of care in avoiding such factors, but in so far as we *have* avoided them, we have reason to suppose that our measures of reliability are also measures of the *validity* of our results.

An example will illustrate the use of the formulæ: An

intelligence examination was given to 64 boys. The average score was 152.4 with a σ_{dis} of 18.4. How reliable is this average or how likely is a new average to fall within certain limits? Applying the formula $\sigma_{av} = \frac{\sigma_{dis}}{\sqrt{n}}$ we find $\sigma_{av} = \frac{18.4}{8} = 2.3$. This signifies that in future experiments under like conditions the chances are 2 out of 3 (for $\pm 1\sigma_{av}$ includes 68.26% of the cases) that the true average lies within the limits $152.4 + 2.3$ and $152.4 - 2.3$ or between 154.7 and 150.1. There is very little probability of the true average falling outside $\pm 3\sigma_{av}$ (refer back to Fig. 6) or outside the limits 159.3 and 145.5. The smaller the σ_{av} the greater the reliability of the average, for the smaller the standard error of the average the smaller the range within which the true average is most likely to be found.

Interpretation of PE_{av} and PE_{med} is similar to that of σ_{av}. The standard error of the average is put to important use in computing the reliability of differences between averages.

The Reliability of a Discovered Difference Between Two Central Tendencies.—Suppose that we repeatedly measure a given phenomenon under one set of experimental conditions and find the average value. Suppose then that we alter some single condition, take a new series of trials and obtain a greater average value. The given condition *appears* to make a *difference* in the phenomenon by increasing it. Or suppose that we measure the intelligence, keenness of vision, or other ability in two groups of persons who differ in age or in sex or in eye color or in some other apparently "single" respect and find *differences* in the averages obtained. How

reliable are such discovered differences? Assuming that the groups are typical (true "random" or "unselected" samples) or are selected from similar populations, how certain are we that the apparent effect of the altered condition is "real" or "significant," or that the apparent superiority of one group will be found to hold in comparing two other similar groups?

Suppose that we discover that the average of distribution B is greater than the average of distribution A, and wish to know how reliable that difference is, that is to say, how likely we are to find the same direction of difference in future series of similar measurements. Evidently our judgments can be based in part upon the amount of the difference discovered, for if our two distributions are at all typical the greater the difference, other things being equal, the less the chance that its direction will be reversed later.

But we must also remember that such a discovered difference depends upon two averages which are themselves somewhat unreliable. In computing its reliability, therefore, we must take into account the reliability of the averages. If the averages are likely to vary little, the difference will vary little; if the averages are likely to vary greatly in future measurements, the difference is also likely to vary greatly. We measure this *tendency of the difference to vary* in terms of values such as the standard error of the difference (σ_D) or the probable error of the difference (PE_D). The formulæ are given below. In the formulæ subscripts A and B indicate the two distributions under consideration.

$$\sigma_D = \sqrt{\sigma^2_{av_A} + \sigma^2_{av_B}}$$

$$PE_D = \sqrt{PE^2_{av_A} + PE^2_{av_B}}$$

Let us call the ratios $\frac{D}{\sigma_D}$ and $\frac{D}{PE_D}$ coefficients of reliability of the difference between averages, for the greater the value of these ratios the greater the reliability of the discovered difference. If this coefficient is as great as 2.7 for $\frac{D}{\sigma_D}$ or 4.0 for $\frac{D}{PE_D}$, and if the measures are typical, the direction of difference discovered (Av_B greater than Av_A, for example) stands nearly 100 chances (99.7) in 100 of being confirmed by later experiments if the assumptions underlying it are justified. If the coefficient of reliability of the difference is less than 2.7 (or 4.0), the probability of such confirmation is less than 100 chances in 100. To determine the actual number of chances in 100 in such cases, reference is made to one of the tables below.

TABLE 11

Shows by Reference to $\frac{D}{\sigma_D}$ the Number of Chances in 100, That the Direction of Difference Discovered Between Two Averages of Two Typical Distributions Will Be Confirmed by Later Experiments

$\frac{D}{\sigma_D}$	.1	.2	.3	.4	.5	.6	.7	.8
Chances in 100	54	58	62	65	69	73	76	79
$\frac{D}{\sigma_D}$	.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6
Chances in 100	82	84	86	88	90	92	93	94
$\frac{D}{\sigma_D}$	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4
Chances in 100	96	96	97	98	98	99	99	99
$\frac{D}{\sigma_D}$	2.5	2.6	2.7	2.8	2.9	3.0		
Chances in 100	99 (99.4)	99 (99.5)	100 (99.7)	100 (99.7)	100 (99.8)	100 (99.9)		

TABLE 12

SHOWS BY REFERENCE TO $\frac{D}{PE_D}$ THE NUMBER OF CHANCES IN 100, THAT THE DIRECTION OF DIFFERENCE DISCOVERED BETWEEN THE AVERAGES OF TWO TYPICAL DISTRIBUTIONS WILL BE CONFIRMED BY LATER EXPERIMENTS

$\frac{D}{PE_D}$	.1	.2	.3	.4	.5	.6	.7	.8
Chances in 100	52	55	58	60	63	66	68	71
$\frac{D}{PE_D}$	.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6
Chances in 100	73	75	77	79	81	83	84	86
$\frac{D}{PE_D}$	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4
Chances in 100	87	89	90	91	92	93	94	94
$\frac{D}{PE_D}$	2.5	2.6	2.7	2.8	2.9	3.0	3.1	3.2
Chances in 100	95	96	97	97	97	98	98	98
$\frac{D}{PE_D}$	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
Chances in 100	99	99	99	99	99	99	100	100
						(99.5)	(99.6)	(99.7)

Although a ratio of $\frac{D}{\sigma_D}$ of 2.7 or above, or $\frac{D}{PE_D}$ or 4.0 or above yield nearly 100 chances in 100 that the discovered difference will be in the same direction in future experiments, actually we never achieve perfect certainty however large the ratio. This, together with the rapid increase in reliability of a discovered difference with increase in the size of the coefficient of reliability of the difference is shown by taking $\frac{D}{PE_D}$ in unit steps as in Table 13.

TABLE 13

CHANCES IN 100 FOR UNIT STEPS

$\frac{D}{PE_D}$	Chances in 100	$\frac{D}{PE_D}$	Chances in 100
1.0	75	5.0	99.98
2.0	91	6.0	99.999
3.0	98	7.0	99.9999
4.0	99.7	8.0	99.999999
		etc.	etc.

For practical purposes any ratio of $\frac{D}{\sigma_D}$ equal to 2.7 or *above*, or of $\frac{D}{PE_D}$ equal to 4.0 or *above*, may be considered to yield approximately *100 chances in 100* that the discovered difference will be in the same direction on future experiments.

Example: Intelligence test scores were obtained for 625 men and 625 women. The averages and σ_{dis} for each were as follows: For boys, $Av = 129.8$, $\sigma_{dis} = 24.5$; for girls, $Av = 127.6$, $\sigma_{dis} = 21.8$. How significant is the difference between averages? For boys $\sigma_{av_A} = \frac{\sigma_{dis}}{\sqrt{n}} = \frac{24.5}{25} = .98.00$; for girls $\sigma_{av_B} = \frac{21.8}{25} = .872$.

$$\sigma_D = \sqrt{\sigma^2{}_{av_A} + \sigma^2{}_{av_B}} = \sqrt{.9604 + .7604} = \sqrt{1.72} = 1.31.$$

$$D = 129.8 - 126.6 = 2.2 \quad \frac{D}{\sigma_D} = \frac{2.2}{1.31} = 1.7.$$

From Table 11, there are 96 chances in 100 that the direction of the difference will be the same in future experiments.

It is customary to state that either a $\frac{D}{\sigma_D}$ of 3.0 or a $\frac{D}{PE_D}$ of 4.0 indicates complete reliability. However, as

pointed out above, there is no such thing as complete reliability. Furthermore, as an example, we cannot state that a $\frac{D}{\sigma_D}$ of 1.4 is unreliable while a $\frac{D}{\sigma_D}$ of 2.0 is reliable. All that we can say is that one is more reliable than the other.

In order to determine mathematically the reliability or "significance" of a discovered difference between two averages, then, the following steps in computation should be taken:

(1) Determine for each series: σ_{dis} or PE_{dis}.

(2) Determine for each series: σ_{av} by dividing σ_{dis} by $\sqrt{n}$; or PE_{av} by dividing PE_{dis} by $\sqrt{n}$.

(3) Determine the measure of variability of the difference by squaring the measures of unreliability of the two averages (σ_{av} or PE_{av}) adding the squares and extracting the square root of the sum.*

(4) Determine D, the difference between the averages.

(5) Divide D by σ_D or PE_D to obtain the coefficient of reliability.

(6) Look up in the appropriate table the chances in 100 corresponding to the coefficient obtained.

Reliability of a Coefficient of Correlation.—The reliability of a coefficient of correlation depends first upon the *magnitude of the coefficient* itself, and secondly upon the *square root of the number of cases.* It is shown by

* When dealing with correlated measures the above formulæ for σ_D and PE_D do not apply. In such a case one of the following formulæ, which take account of the correlation, should be used:

$$\sigma_D = \sqrt{\sigma^2_{av_A} + \sigma^2_{av_B} - 2r\sigma_{av_A}\sigma_{av_B}}$$

$PE_D = \sqrt{PE^2_{av_A} + PE^2_{av_B} - 2r\,PE_{av_A}PE_{av_B}}$ where r is the coefficient of correlation. (Deduce r from ρ by $r = 1.0233\rho$.) See Garrett, p. 287.

the probable error of ρ and is computed by means of the formula

$$PE_{\rho} = \frac{.706\ (1 - \rho^2)}{\sqrt{n}}$$

A coefficient of $+ .64$ with a PE of .04 is written as $+ .64 \pm .04$ and means that, in a future experiment under like conditions, the chances are 1 to 1 (for ± 1 PE includes 50 per cent of the cases) that the obtained ρ will fall within the limits of $.64 \pm .04$, or .60 and .68. The chances are 99.3 in 100 that such a coefficient will fall within the limits of ± 4 PE (see normal curve) or between .48 and .80. For any given coefficient, the smaller the PE the more dependable is the coefficient.

For fine distinctions such as predicting a person's rank in an intelligence test given to a selected superior group, the reliability of ρ should be high. For grosser distinctions the reliability may be less and still have the coefficient of practical use for predictive purposes. That is, the reliability of ρ in a practical situation depends partly on the use to which the coefficient is put.

The common statement is that ρ should be at least 4 times its PE before the coefficient is dependable enough to indicate that there is some correlation present. However, we have no right to state that a $\frac{\rho}{PE_{\rho}}$ of 4.0 indicates a reliable coefficient and a $\frac{\rho}{PE_{\rho}}$ of 3.7 an unreliable one. All that we are justified in saying is that one is more reliable or dependable than the other. The larger the ratio between a coefficient of correlation and its PE the more reliable is the coefficient, other things being equal. This is well shown in Table 13. To use this,

substitute the term $\frac{\rho}{PE_\rho}$ for $\frac{D}{PE_D}$ in the table. Then for any value of ratio $\frac{\rho}{PE_\rho}$ up to 8.0, note the corresponding chances in 100 that the direction of relationship (plus or minus) will be maintained in future experiments under like conditions. For example, the correlation coefficient of $+ .60 \pm .20$ between 2 series of variables indicates that there are 98 chances in 100 that future measurements will show a positive correlation between the 2 series. Similarly the coefficient $- .64 \pm .08$ indicates that there are 99.999999 in 100 that future measurements (under like conditions) will show a negative correlation. When the ratio $\frac{\rho}{PE_\rho}$ is greater than 8.0 the chances of confirming the direction of the discovered relationship on future experiments are *very high* but, of course, absolute certainty is never achieved. Always interpret reliability of ρ with reference to the situation in which it is used.

During the rest of this course obtain a PE for each ρ computed. Use Table 16 of Appendix to get PE_ρ.

Comparative Variability.—When two distributions are compared with regard to variability, we need to remember that the respective measures of variability of the two are distances from different measures of central tendency. If the two measures of central tendency lie close together, direct comparison of their respective AD's, σ's or PE's (of the distribution) may not be misleading. But if the averages are not approximately the same, direct comparison of the measures of variability is apt to be misleading. If, for example, the average weight of American men was found to be 160 pounds and the AD from this average was discovered to be only

8 pounds, we should probably say at once that men vary relatively little in weight. But if the AD from the average weight (say 16 pounds) of one-year-old babies was found to be 8 pounds we should say that the weights of one-year-old babies vary greatly. The absolute measures of variation (the AD's) for the two distributions are the same (8 pounds) but their relative variabilities are clearly not. In comparing variabilities of two distributions we ought, then, in strictness to take into consideration the measures of central tendencies from which the variations are computed. Relative variability, therefore, is usually stated not simply as an amount, but as a percentage of the measure of central tendency This percentage is called the coefficient of variability and is designated by V. One of the following formulæ may be used to compute it, depending of course, upon the measure of variability employed:

$$V = \frac{100\ AD}{Av} \text{ or } V = \frac{100\ \sigma}{Av} \text{ or } V = \frac{100\ PE}{Av}.$$

Thus the coefficient of variability for the weights of men in the supposed case is $\frac{8 \times 100}{160} = 5\%$, and the coefficient of variability for the weights of babies is $\frac{8 \times 100}{16} = 50\%$. Therefore, the weights of men are $\frac{5 \times 100}{50} = 10\%$ as variable as that of the babies.

If, for example in an arithmetic test, a class of boys obtained an average score of 60 with an AD of 12; a class of girls an average score of 64 with an AD of 16, then the coefficient of variability for boys is $\frac{12 \times 100}{60} = 20\%$; that for girls $\frac{16 \times 100}{64} = 25\%$. The boys are $\frac{20 \times 100}{25} = 80\%$ as variable as the girls in this test.

"The coefficient of variation is especially useful in those problems in which the variability of the group under different conditions is the factor studied." *

Exercises.—Keep figures correct to the second decimal place.

(1) Judging from the coefficients of variability (use σ) for exercise (4), Chapter III, which were relatively more variable in intelligence—boys or girls? How many times more variable?

(2) Judging from the coefficients of variability (use AD) for exercise (9), Chapter III, which showed relatively more variability in performance—men or women? How many times more variable?

(3) In first trial in an experiment in mapping the sensory endings for warmth on a given area of the skin students report: 1, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 7, 8, 8, 8, 8, 9, 9, 9, 9, 9, 9, 10, 11, 11, 12, 12, 12, 13, 13, 14, 14, 14, 15, 16, 16.

On second trial with warm spots students report: 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 9, 9, 9, 9, 10, 10, 10, 10, 11, 12, 12, 12, 12, 15, 16.

Compute averages. Then compute $\sigma_{av} = \frac{\sigma_{dis}}{\sqrt{n}}$ for each series. Obtain D between averages. Compute σ_D and $\frac{D}{\sigma_D}$. Look up chances in 100 in Table 11.

Judging from these data, how great is the probability that students will find more warm spots on the first trial than on the second trial in future experiments?

(4) Two tests were given to 169 officers and to 196 candidates for officers' commissions.

* See Garrett, p. 42.

Test A gave the following scores:

for officers . . . Av = 32.6 $PE_{dis} = 4.2$

for candidates Av = 30.5 $PE_{dis} = 2.5$

Test B gave the following scores:

for officers . . . Av = 29.7 $PE_{dis} = 5.2$

for candidates Av = 28.0 $PE_{dis} = 4.6$

If we may judge from these particular groups, how many chances in 100 are there that officers as a class will make better scores than candidates in test A? How many chances that they will in test B? Which, then, is probably the better test of military competence—A or B?

(5) 100 adult feeble-minded and 225 adult normal persons were given tests A and B.

Test A gave the following scores:

for the feeble-minded. . Av. = 9.9 $PE_{dis} = 4.4$

for the normal persons Av. = 23.49 $PE_{dis} = 5.8$

Test B gave the following scores:

for the feeble-minded. . Av. = 2.5 $PE_{dis} = 2.8$

for the normal persons Av. = 15.64 $PE_{dis} = 7.1$

Judging by the coefficient of reliability of the difference, which test distinguishes better the feeble minded from the normals? Are we certain from the above computations that all feeble-minded persons will make poorer scores than any normal person on these tests? That the average score of a typical group of feeble-minded will be less than the average score of a typical group of normal adults?

References

H. E. Garrett: *Statistics in Psychology and Education*, 1926; A. S. Otis: *Statistical Method in Educational Measurement*, 1925; T. L. Kelley: *Statistical Method*,

1924. On $\frac{D}{PE_D}$ as a measure of reliability, see E. G. Boring, *American Journal of Psychology*, 1916, Vol. 27, pp. 315-319; 1917, Vol. 28, pp. 454-459; 1920, Vol. 31, pp. 1-33; *Psychological Bulletin*, 1919, Vol. 16, pp. 335-338.

CHAPTER XXVIII

Experiment No. 22

A GROUP INTELLIGENCE EXAMINATION

(The Army Alpha Test *)

As the preceding discussion has indicated the reliability of an intelligence test is often determined by correlating score with retest score or correlating score on one form with score on another form. The validity of an intelligence test is commonly obtained by correlating scores obtained in it with estimates or other outside criteria of intelligence. In devising a test or scale of tests for the measurement of intelligence, the experimenter usually begins with a general concept or a general analysis of the nature of intelligence, invents or uses tests whose performance appears to him to involve the ability or part-abilities concerned, gives such tests to groups whose ability is known, compares the averages of differing groups and correlates the scores with the criteria, rejects the tests which discriminate or correlate least well, and finally relates the remaining tests to norms or standards of performance. Since in actual

* Any other group test which has two equivalent forms, takes about 40 minutes of examining time, and consists of 4 to 10 parts, may be used in place of Alpha. Two such tests are: The Otis Group Intelligence Scale, Advanced Examination, by A. S. Otis, published by World Book Co., Yonkers-on-Hudson, N. Y.; and Psychological Examination for High School Grades and College Freshmen, by L. L. Thurstone, published by American Council on Education, 26 Jackson Place, Washington, D. C.

practice the criteria are variable, the groups never perfectly unselected, and the correlations never perfect, no test or scale is ever a perfectly reliable or perfectly valid one. After a certain degree of validity is attained, even long and persistent trial and modification with all due care to minimize the sources of error, increase our assurance only gradually. The standardization of a test is a work which never ceases until a whole population has been examined.

The Army Alpha Test is one of the best known and best standardized tests for adults among the great number at present existing. It is composed of eight tests chosen for both practical and theoretical reasons by a committee of experts, after extended trial and searching examination of all available evidence regarding their value and reliability. In order to minimize the possibility of coaching several approximately equivalent forms of the test are available. These have been given to some million persons, and norms based upon a random sample of over 100,000 men in the U. S. Army may be consulted. The test has shown itself to be valuable not only in the selection of men for the purposes for which it was devised, but to have a much wider usefulness and a remarkably high general validity. An account of the history, development, use, and results of the test, covering nearly 900 octavo pages, has been published by the National Academy of Science.

The purpose of this experiment is: (1) to determine the influence of practice and of sex on intelligence test scores; (2) to compute the reliability and validity of the Alpha Test for a group of college students; and (3) to illustrate the use of group intelligence tests.

Apparatus.—Timer or stop watch. Manual of Instructions ("Army Mental Tests" by Yoakum and Yerkes). Test blanks for two forms of Alpha Test.

Method.—The instructor acts as experimenter and the class as subjects. One form of Alpha is given to the whole class according to directions in the manual. Instruct the students to write "Form A, taken first" as well as their name and sex on the top of the first page of the test. Collect all examination blanks and allow a rest period of about ten minutes. Now give the alternative form to each subject. Have the students label this "Form B, taken second." First one form and then the other is scored by the class by reference to the correct answers read by the instructor. Have the students exchange blanks with their seat neighbors for scoring. The data for the men are then listed in one tabular form and that for women in another, either on the blackboard, or on a sheet prepared by the instructor. The class material is then copied from the blackboard or from dictation for use in computations by individuals.

Arrangement of columns in each table:

Column 1—Total score for Form A (first examination).

Column 2—Total score for Form B (second examination).

Column 3—Score in Test 2 of Form A.

Column 4—Score in Test 8 of Form A.

Column 5—Class grades given the students on the mid-quarter examination or for the first half of the course (or honor-point average for the previous quarter).

Results and Discussion.

(1) How much on the average do students gain in score in the second examination (use combined data; men and women)? Judging from these data, is there a significant gain in score on the second examination? Base your answer on the computation of chances in 100 that the average of all scores in one form will be greater than the average of all scores in the other. Interpret clearly the comparison thus made. Use $\sigma_{av} = \frac{\sigma_{dis}}{\sqrt{n}}$ in computations here and in question (2).

(2) Does the Alpha test appear to be significantly harder for men or for women? Compare the average score for men with the average score for women in Form A of the examination and find the chances in 100 indicated by these data. Explain.

(3) What percentage of the class show no improvement in the second examination? State the significance of this finding.

(4) Correlate total scores in Form A with total scores in Form B and obtain the PE of ρ. What does this coefficient show concerning the reliability of the Alpha examination? The reliability of the Alpha test, as used in the army, was $+ .95$. How does this compare with your reliability coefficient? What factors different from those in the army, are present in our experiment and consequently might be expected to change the reliability of Alpha scores in this group?

(5) Using the data from Form A, correlate scores in test 2 with total scores for the rest of the test (i.e., total score minus score on test 2). Represent the correlation graphically. Assume that total score

(the rest of the test) measures intelligence accurately. Then by reference to the graph show that some individuals would be badly judged, if it was tried to measure their intelligence by test 2 alone. If test 2 correlated very highly with the rest of the examination it would be superfluous. Explain why.

(6) What is the correlation between the total score of Form A and class grades (use either class grades or honor point averages for previous quarter)? If we assume that total score measures intelligence accurately, then how valid are Alpha scores as a measure of aptitude in laboratory psychology (or aptitude in college)? Do you consider intelligence the only factor entering into grades? Explain. List other factors which you think influence school grades.

(7) To what practical uses can intelligence (college aptitude) test scores for college students be put? (See Pintner, first reference, pp. 265-279.)

Summary and Conclusions.

References

R. Pintner: *Intelligence Testing,* 1923, esp. pp. 265-282; R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 23-69; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 453-478; C. S. Yoakum and R. M. Yerkes: *Army Mental Tests,* 1919; *Memoirs of the National Academy of Science,* Vol. 15, 1921; F. A. C. Perrin and D. B. Klein: *Psychology,* 1926, pp. 313-333; D. Starch: *Educational Psychology,* 1927 (Rev. Ed.), pp. 107-124; A. M. Jordan: *Educational Psychology,* 1928, pp. 343-380; R. Pintner: *Educational Psychology,* 1929, pp. 89-115.

CHAPTER XXIX

Experiment No. 23

THE MEASUREMENT OF MUSICAL TALENT

Discrepancies between general intellectual capacity (intelligence test scores) and actual achievement in many particular tasks suggest the existence of many aptitudes not measured by general intelligence tests. Special aptitudes of various kinds are now being measured objectively by standardized tests and scales. In the field of music we have objective measurements of musical aptitudes, musical accomplishment, and musical appreciation. Continued progress in this direction assures future changes in both pedagogy and vocational guidance in the field of music. The Seashore Measures of Musical Talent furnish a good example of special aptitude tests in the musical field. After careful experimentation in the psychological laboratories at the University of Iowa, C. E. Seashore and his colleagues constructed a number of psychological tests of musical talent and issued them in the form of phonograph records. These tests are based upon an analysis of the sensory equipment which seems to be indispensable for a musician. Six tests were devised:

(1) *Sense of Pitch.*—This measures pitch discrimination in terms of the smallest difference in pitch that the subject can recognize. The pitch difference in the test ranges in difficulty from 30 vibration intervals (easy) to ½ vibration intervals (difficult). The stimuli repro-

duced by the record are produced by tuning forks whose tones are amplified by resonators.

(2) *Intensity Discrimination.*—Sense of intensity is measured in terms of the smallest difference in loudness of sounds perceptible to the subject. The range in loudness varies from 5 units (easy) to 1. The original sounds are produced by an apparatus called the audiometer.

(3) *Time Discrimination.*—This measures the least perceptible difference in time that the subject can hear. The time intervals vary from 1.00 to 1.20 sec. and yield differences in time which vary from .02 to .20 sec.

(4) *Consonance Discrimination.*—Here the subject judges consonance on the basis of smoothness, blending, and purity of the stimuli.

(5) *The Tonal Memory Test* measures the subject's memory span for tones unrelated in melody.

(6) *Sense of Rhythm* tests the subject's capacity to perceive differences in rhythmic patterns on the basis of time and intensity.

Although the tests are not entirely free from imperfections they nevertheless yield a fairly reliable measure of certain basic psychological capacities. A recent restandardization shows little change from the first norms. Seashore's coefficients of reliability * on an adult group of 200 persons follow:

Pitch	.809	Time	.681	Memory	.919
Intensity	.750	Consonance	.714	Rhythm	.681

On the consonance and rhythm tests with children the reliabilities are somewhat lower than the above. Other

* Cited by permission of Professor C. E. Seashore. New norms and reliabilities will soon be published in a forthcoming monograph. See the monograph for method of computing reliability.

investigators report somewhat lower reliabilities than Seashore (see references). When validated against teachers' ratings and expert musicians' ratings for musical ability the validity coefficients vary from .15 ± .06 to .47 ± .12, the latter being for memory and rhythm. These higher coefficients compare favorably with validities of group intelligence tests as validated against school grades.

The results of these musical aptitude tests are of considerable practical importance. As Seashore points out, the negative results of the tests are more valuable than the positive. That is, if an individual is below a certain score it would be impossible for him to achieve much in music. For example, Seashore * suggests the following use of scores on the pitch discrimination test for advising subjects concerning musical training. Those receiving the highest 10 per cent, stimulate enthusiastically; next 20 per cent, encourage freely; next 40 per cent, encourage; next 20 per cent, question; next 10 per cent, discourage.

These six measures of musical ability do not constitute a complete survey of musical talent as Seashore has clearly shown by his analysis of a musical mind. However, they are specific measures of six important basic capacities. Since these are specific and fundamental traits it is impossible to combine the scores on the six tests into a score of musical talent. A concise picture of an individual's musical talent in relation to the norms is given by plotting his proficiency profile or "psychograph."

The Proficiency Profile.—The abilities of no person, as determined by standardized measuring devices, are

* Measurement of Musical Talent, pp. 67-68.

uniform even when experimental conditions are very rigidly controlled. Accomplishment in school subjects shows this characteristic tendency toward lack of uniformity. Scores on achievement tests demonstrate that no student exhibits exactly the same degree of ability in all school subjects. This variability in achievement is very clearly displayed in "proficiency profiles" or "psychographs." Figure 18 is the proficiency profile of an elementary school student. The heavy horizontal line represents the median or 50th percentile ability in each test. (A percentile is a point on the scale of distribution. Thus the 90th percentile is the point above which lie $1/10$ and below which lie $9/10$ of the measures. Hence Q_1 is the 25th percentile and Q_3 the 75th percentile.) The degree of ability or percentile rank of pupil M for each school subject is represented on the graph by a short horizontal line drawn across the rectangle above the name of the test. Examination of the profile of this student shows superior performance in arithmetic, science and geography, but poor performance in writing, composition and spelling. In drawing, reading and history he ranks somewhat above median. In general subject M exhibits rather large variability in school achievement.

Psychographs may also be constructed to represent abilities measured by other tests such as the Seashore tests of musical talent to be employed in this experiment. To have much significance a proficiency profile must be based on measures of relatively high reliability.*

The Graphic Rating Scale.—A method useful for measuring the value of various kinds of stimuli as well

* Not all of the Seashore musical tests are highly reliable measures and proficiency profiles constructed from their scores should be interpreted with this limitation in mind.

as certain capacities and traits is *rating on an absolute scale.* The graphic rating scale is one of the more usable rating scales and is perhaps more interesting than others. Graphic rating involves the name or definition of the

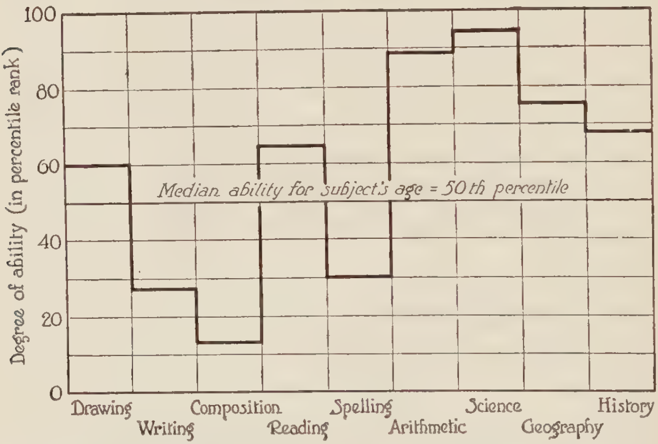


Fig. 18. Proficiency profile showing percentile rank of pupil M in elementary school subjects.

trait to be rated, or both, followed by a straight line a few inches long representing the distribution of the trait from maximum to minimum. Along and underneath the line are placed descriptive adjectives for the guidance of the rater. The descriptive phrases are so arranged that those at one end of the line indicate a high degree of possession of the trait, those at the other end, a low degree. Intermediate degrees of possession range between the extremes. Guided by these adjectives the rater checks at some point along the line, not necessarily directly underneath a descriptive phrase. For example, suppose you are considering the neatness and accuracy of a stenographer's work and her ability to

maintain high workmanship in these respects. The place on the following scale which best describes your estimate of the trait is to be checked:

Neatness and accuracy	Makes many errors	Careless	Mediocre	Neat and accurate	Does excellent work

The distance from the left-hand end of the scale to the place checked would represent the score for this trait. Sometimes the scale is scored on the basis of a certain number of units, as 10, rather than in millimeters or sixteenths of an inch. This kind of measurement can be applied to various personality and character traits, and can also be used to measure such things as the affective value of musical intervals, self-judgments of musical ability, etc.

The purpose of this experiment is to illustrate objective aptitude tests of musical talent, to compare musical scores of a college class with adult norms, to study the relation between self-ratings of musical ability and scores on the objective tests, and to determine the relation between training and test scores.

Apparatus.—Phonograph, the six Seashore double-disc records containing measures of musical talent, manual of directions for giving the Seashore tests, musical rating blank (see *Notes for Instructors*).

Method.—Have each student rate himself on musical training and musical talent according to the musical rating blank.

In giving the Seashore tests the instructor acts as experimenter and the class as subjects. See that the class is sitting as near the phonograph as possible. Set the

phonograph for the proper speed and give the six tests as directed in the manual. Furnish students with mimeographed blank forms for recording responses. After each test (pitch, intensity, etc.), have the members of the class score their papers while the instructor reads off the correct responses. This is done by drawing a straight line through the wrong answers. Then count the number of mistakes in the whole record and subtract this from the total number of trials. This will give the number of right answers. Reduce this to a percentage by dividing the number of correct answers by the total number of trials. Place this score at the top of the record. Designate each member of the class by a number and, as he reads this score, translate it into percentile ranking (see tables in Seashore Manual) and tabulate this on the board for each of the six tests. Ask the students to copy these data for use in their computations.

Results and Discussion.

(1) Obtain the median percentile rank for the class in each of the six Seashore tests. What explanation would seem to you to be adequate for the difference found in comparing this class with Seashore's norms (medians of Seashore's norms are the 50th percentiles)?

(2) What are the coefficients of correlation between pitch and the other five tests? Are these correlations all rather high or low? In a test of musical talent is it desirable to have low or high correlations between the various parts of the test, i.e., between pitch and intensity, etc.? Explain. (The instructor should discuss intercorrelations between parts of a test.)

(3) What is the correlation between musical training score and musical talent score (self-ratings). State carefully what this correlation tends to show. What is the PE of ρ? Is the obtained coefficient reliable?

(4) Correlate musical training score with pitch and also with memory. Do the same for musical talent scores (self-ratings) and pitch and memory. Assume that these coefficients of correlation between musical self-ratings and these two objective tests of musical ability are representative of what would appear from correlating the rating scores with scores on the other objective tests of musical talent. What do these correlation coefficients tend to show?

(5) State carefully and at length what you consider to be the reasons for the difference between the correlation in question 4 and those in question 5.

(6) According to experimental results, what (briefly) are the effects of heredity and environment on musicality? (See Farnsworth, pp. 233-246.)

(7) The Seashore tests of musical talent are classified under what general type of tests? Why? (See Gates, pp. 484-487).

(8) Construct a proficiency profile to show your ability on the 6 Seashore tests in relation to the median of the norms. Interpret in detail the proficiency profile.

(9) Describe a method of computing reliability on the pitch discrimination test.

Summary and Conclusions.

References

A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 484-489; F. A. C. Perrin and D. B. Klein: *Psychology,* 1926, pp. 334-335; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 318-320; P. R. Farnsworth: *The Effects of Nature and Nurture on Musicality,* Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, Part II, pp. 233-246; C. E. Seashore: *The Psychology of Musical Talent,* 1919; J. Kwalwasser: "Tests and Measurements in Music," *Psychol. Bull.,* 1928, Vol. 25, pp. 284-301; H. M. Stanton: *Psychological Tests of Musical Talent,* Eastman School of Music, 1925; F. M. Brennen: "The Relation Between Musical Capacity and Performance," *Psychol. Monog.,* 1927, Vol. 36, pp. 200-248; G. M. Ruch and G. D. Stoddard: *Tests and Measurements in High School Instruction,* 1927, pp. 195 ff.; A. W. Brown: "The Reliability and Validity of the Seashore Tests of Musical Talent," *Jour. Appl. Psychol.,* 1928, Vol. 12, pp. 468-476.

For rating scales, see H. E. Burtt: *Principles of Employment Psychology,* 1926, pp. 317-370.

CHAPTER XXX

Experiment No. 24

THE MEASUREMENT OF READING ABILITY

Reading ability has an important bearing upon both recreational reading and study. This aptitude determines the amount of material covered within a certain period as well as the degree of comprehension. Reading is the most common method employed to acquire knowledge in the various fields of study. Recently the most progressive schools are devoting increased attention to the development of efficient habits of reading and study. Because silent reading is used almost exclusively in study and recreational reading, it receives more consideration than oral reading in teaching programs, especially after the first years of school life. In these programs there is a rather general acceptance of significant practical contributions of research in the field of reading.

Three important factors involved in reading ability are: (1) speed of reading, (2) vocabulary knowledge, and (3) comprehension. Tests of high reliability are now available for the three factors cited. The reliabilities for the tests to be used in this experiment are as follows: Speed of reading, Form 1 *vs.* Form 2, $.80 \pm .03$; vocabulary knowledge, test *vs.* retest, $.91 \pm 01$; comprehension, test *vs.* retest, $.78 \pm .03$. Ability in each of these phases of reading is usually related in some degree to ability in the others. Comprehension and vocabulary

knowledge also seem to be fairly closely related to intelligence test scores. Achievements on these reading tests show a wide range of individual differences for all three measures of reading ability.

Growth in reading ability reveals certain developmental characteristics. Rate and comprehension in reading increase rather constantly from the early grades up to high school and thereafter more slowly or not at all. Both oral and silent reading improve markedly with careful training for a few weeks. How permanent such improvement is, however, has not been shown.

Eye movement habits seem to be closely related to efficiency in reading. As anyone may note by watching another read, the eyes move by jerks or *saccades* rather than in a continuous sweep during the process of reading. At the end of each jerk or movement comes a short pause. Since the time taken for movements is comparatively very short, a large proportion of the total reading time is devoted to fixational pauses (about 95 per cent). It has been demonstrated that there is no clear vision during the saccadic eye movements occurring in reading. The use of relatively few pauses per line is an indication of efficient reading. Ordinarily the efficient reader shows a steady progression of eye movements from left to right along the line of print. Movements occurring within the line and in the opposite direction, right to left, are termed regressions. For the most part these regressions are due to poor eye movement habits and their frequent occurrence is a mark of inefficient reading. The return movement from the end of one line to the beginning of the next is called a return sweep. Saccadic eye movements are also used in examining pictures, landscapes, or other stationary fields of view.

Investigations of eye movement habits have been made by analyzing photographic records of eye movements in reading. Some of the problems studied are: the development of eye movement habits with progress in reading ability; comparison of oral and silent reading; reading foreign languages; influence of reading attitude on eye movements; the reading of arithmetical problems; reading formulæ; the diagnosing of reading difficulties.

An accurate study of eye movements in reading requires the use of some mechanical means of recording them, such as the elaborate photographic method devised by Dodge and used at various universities in reading investigations. The main characteristics of eye movements may be observed, however, by holding a mirror beside the reading text so that the experimenter, sitting beside the reader, can observe and count the eye movements. This method is crude but yields an approximate count of the movements and demonstrates very well the nature of eye movements in reading.

The purpose of this experiment is: (1) to investigate individual differences in speed of reading, comprehension in reading, and word knowledge; (2) to note the correspondence between scores in these measures, and between intelligence and reading ability; and (3) to investigate the nature of eye movements in reading.

Apparatus.—Timer or stop watch. Manuals of instructions. Test blanks * for the Minnesota Speed of Reading Test and Minnesota Reading Examination which includes vocabulary and comprehension tests. (See Appendix.) Mirror and selections for reading. (See Appendix.)

* Any other reading test or tests which contain measures of rate, comprehension, and vocabulary may be used instead of the ones given here. See *Notes for Instructors* for suggestions.

Method.

Part 1. *Reading Tests.*—The instructor acts as experimenter and the class as subjects. First a single form of the Speed of Reading Test is given to the whole class according to directions in the manual. Have the students exchange blanks and then score the tests by reference to the correct answers read by the instructor. Give the vocabulary and comprehension tests to the whole class according to directions in the manual but use 30 minutes instead of 40 for comprehension test. The students again exchange blanks and score the tests as the instructor reads the correct answers. The data from the whole class plus the Alpha intelligence test score of Form A from Experiment No. 22 are then listed in tabular form on the blackboard or on a sheet prepared by the instructor. The class material is copied from the blackboard or from dictation for use in computations by individuals.

Arrangement of columns in the table:

Column 1—Score in speed of reading test.
Column 2—Score in vocabulary test.
Column 3—Score in comprehension test.
Column 4—Combined score of vocabulary and comprehension tests.
Column 5—Total score for Form A of Army Alpha intelligence examination (from Experiment No. 22).

Part 2. *Eye Movements in Reading.*—This part of the experiment, which is to determine the nature of eye movements in reading, is to be done by students working in pairs. Obtain a mirror from the instructor.

Instruct S: *"You are to look into the mirror and fixate your vision on the image of your right eye; now*

change your fixation to the left eye and try to detect the movement of the eyes during the shift from one position to the other. Do not mistake slight movements of the eyelids for eye movements. Try this several times and report your success to E."

E records the report of S. Failure to see the eyes in motion confirms the fact that there is no clear vision during eye movements of this type.

Place the mirror, with its far (from E) end resting on a book, at the right side of the text and in such a position that E, who sits beside S, is able to see clearly the image of the S's right eye. Now have S read the first few lines of the introduction to this experiment while E watches the images of his eyes and notes the quick, jerky movements followed by pauses. S should incline the head forward a little so that his eyelids are well raised. Have S fixate the point of E's pencil while he moves it slowly to and fro and at the same time watches S's eyes. Describe these eye movements. They are termed pursuit movements and permit clear vision during their execution.

Turn to the practice exercise in the Appendix and instruct S: "*Fixate your vision at the very beginning of the first line in the practice exercise. When I say 'go' begin and read through to the end of the selection. After finishing the last line look over to me quickly and say 'done.' You are to read the selection for the meaning and at your normal rate. After the reading you are to state in your own words the contents of the selection.*"

E is to count the total number of eye movements, including the return sweeps for the whole selection. This is best done by watching the edge of the iris (colored part of the eye) in relation to the eyelid. As there is

a fixation following every movement, there will be as many fixations in reading as there are movements with the exception of the last movement in turning to E at the end of the reading. Have S tell in his own words what he read.

Divide the total number of fixations by the number of lines to obtain the number of fixational pauses per line. Repeat this exercise *three or four times* until E is confident that he is not missing any of the eye movements.

After adequate practice on the preliminary exercise, have S read exercise A once in the same manner and obtain the number of pauses per line. S states the contents of the passage as before. Now have S read exercise B while E tries to detect regressive movements, i.e., movements from right to left within the line. Try to count these regressive movements and note, if possible, whether they all occur within the line or whether some occur at the beginning. Those which happen at the beginning occur because the return sweep to the next line is inaccurate and a readjustment is necessary to perceive the first words in the line. Record your findings.

Results and Discussion.

Part 1.

(1) Compute the class medians for each of the 3 reading tests; speed, vocabulary and comprehension. Compare your individual score with the class median on each test. Before taking these tests did you consider yourself an efficient or an inefficient reader? How do these beliefs agree with your reading scores?

(2) Compute the number of pages of material (320 words to the page), equal in difficulty to that in the speed of reading test, that the slowest reader and

the fastest reader in your class could read in a ten-hour period. Each item on the test contains approximately 50 words and 6 minutes were taken for the examination. Compare the scores thus computed for the two students and state the practical significance of your findings.

(3) Correlate speed of reading scores with scores on comprehension test. Explain the significance of the obtained coefficient. A certain textbook states that there is a fairly direct relationship between rate of reading and comprehension. Discuss this statement in the light of your results.

(4) Correlate speed of reading scores with scores on vocabulary; scores on vocabulary with scores on comprehension. Interpret the coefficients obtained. Compare coefficients and try to explain the difference.

(5) Correlate the combined vocabulary and comprehension scores with scores on the Alpha intelligence examination. Interpret carefully the obtained coefficient.

Part 2.

(6) Summarize the results obtained when S attempted to see his own eyes move by watching the image in the mirror. State fully the implications of these results (see Brooks, p. 38). Compare *saccadic* and *pursuit* eye movements.

(7) Describe briefly the motor behavior of the eye in reading, discussing movements, pauses, and regressions (see Brooks, pp. 36-38).

(8) Obtain the average number of pauses per line and average number of regressions per line for your S in reading exercise A. Compare the average num-

ber of pauses per line obtained for S with 6.5, the average obtained for 45 subjects in reading lines 90 mm. in length with photographic recording of movements. Explain differences.

(9) How widely do readers vary in average number of pauses per line and in pause duration (see Brooks, pp. 38-39)?

Summary and Conclusions.

References

F. D. Brooks: *The Applied Psychology of Reading,* 1926, pp. 36-45, 85-130; F. N. Freeman: *Experimental Education,* 1916, pp. 95-109; C. R. Stone: *Silent and Oral Reading,* Boston, 1926 (Rev. Ed.), pp. 6-32; M. A. Tinker: "Legibility and Eye Movement in Reading," *Psychol. Bull.,* 1927, Vol. 24, pp. 621-639; H. G. Wheat: *The Teaching of Reading,* 1923, pp. 50-63, 120-142; W. A. Smith: *The Reading Process,* 1922, pp. 108-129, 152-178, 226-261.

CHAPTER XXXI

Experiment No. 25

THE ATTENTION VALUE OF ADVERTISEMENTS

From the introspective point of view, attention consists in the fact that a mental process may have different degrees of clearness or prominence in consciousness, or that at any given time one or more processes occupy the "focus" and the others the "margin" of consciousness. To pay attention or to give attention means that certain of the mental processes simultaneously present become clear.

From the point of view of behavior, attention is a form of set, a tendency to react selectively and determinatively, to one or a few of the many stimuli which simultaneously affect the neural mechanism. To pay or to give attention means selective reaction, the facilitation of one type of response and the inhibition of other types.

Of the possibly effective stimuli, some "naturally" condition attention. They appeal to native reaction tendencies, neural tendencies present in young children as well as in adults. Such types of stimuli are called the primary conditioning factors of attention. (1) *Intense* things, (2) *large* things, (3) *repeated* things, (4) *sudden* things, (5) *isolated* things, (6) *moving* things, (7) *novel* things belong in this group.

But natural interest or reaction tendencies are modified by experience, and interest in primarily uninterest-

ing things develops. Anything vividly and repeatedly or recently "associated" (connected) with an originally interesting matter takes on an interest and suggestiveness of its own, and henceforth tends by itself to attract and to hold attention. The direction of our attention at any given moment is therefore determined by (8) *habitual* interests (secondary or derived conditions). And once attention is caught, or attentive set established, it tends to persist, so that (9) *congruence* with a present idea or topic of thought, (10) *congruence* with a recent instruction, question, or suggestion, (11) *congruence* with a native reaction tendency or an habitual interest, may also be regarded as secondary conditioning factors. The list of factors, designated by the numbered and italicized words in the two foregoing paragraphs, is by no means a complete or a logically systematic one. Some of the classes mentioned evidently overlap, some are designated indefinitely, and many might be still further subdivided. They are offered as a starting point for the more definite and detailed assignment of factors called for in the experiment which follows.

Advertisements furnish us convenient and rich material for the exhibition and study of such conditions and of their mutual influences. The value of an advertisement does not, of course, depend simply upon its tendency to catch attention or to hold interest, important and primary as this may be. In order to be maximally effective it should also be designed to please, to be remembered, to arouse good-will, to create belief, to persuade, to arouse desire and demand, and ultimately to lead to the action (purchase, employment, etc.), desired by its author. Its design, its appropriate medium, and its economic value also depend upon the character of the public

to be reached, the space available, the cost of insertion or construction, and other factors too numerous to mention.

No perfectly accurate method for measuring the attention value of advertisements has been devised, but properly controlled estimates or judgments by properly selected groups of persons have proven rather more consistent, and rather more valuable when checked against practical results than might have been expected.

The order of merit method offers a partial but extremely simple means of control. The procedure is the following: The stimuli are presented simultaneously to S in a haphazard order. S proceeds systematically to rank them so as to obtain an order of merit or order of excellence for some single characteristic of the series. Thus he may be asked to confine himself first to the question: Which advertisement catches my attention first, or most strongly catches my attention? Which next most strongly, etc? Or S may be similarly set to estimate almost any other definitely formulable characteristic or quality of a series of any sort whatever, provided that he can confine himself closely to that quality and can pass fair and reasonably consistent judgments concerning it. If, however, successive instead of simultaneous presentation of stimuli is necessary, if absolute measures instead of simple relative estimates are required, or even if more detailed relative measures are desired, the method needs to be supplanted by others which give stricter control.

The purpose of this experiment is: (1) to determine by the order of merit method the value of certain advertisements for catching attention; and (2) to determine by the same method the value of the same advertisements

for holding interest. The results in (1) and (2) will be analysed to determine factors involved in attention value and interest in advertisements and to determine sex differences.

Apparatus.—Each student is required to bring to class the current copy of the *Saturday Evening Post,* or a similar magazine containing numerous advertisements.

Method.—All members of the class will serve as subjects, recording their own results. The instructor will designate by page numbers the 15 (or 10 or 12 as the instructor desires) advertisements to be considered. The students remove these pages and lay out the advertisements in order of page number upon a long shelf or table. Each S must do his ranking independently.

Part 1. *Value for Catching Attention.*—Imagine yourself as glancing casually over the pages of the magazine, looking at the advertisements. Try to be your natural self and not to judge as you think you somehow ought to, or as you think someone else would judge. Do not be critical and do not stop to *study* any particular advertisement. With this attitude consider the first two advertisements at the left. Neglect the others for the time being. Which one of the two most definitely and most strongly attracts or catches your attention as you glance at them. Put it at the left in position No. 1, and the other in position No. 2 to the right of No. 1.

Now take any other advertisement, compare it with the advertisement in position No. 1. If it catches your attention more definitely and strongly than No. 1, place it to the left of No. 1. If, however, this new advertisement has less attention value than that in position No. 1, compare it with the advertisement in position No. 2. If this comparison gives it greater attention value than

No. 2, place it between No. 1 and No. 2. Otherwise place it in position No. 3 at the right of No. 2. Continue this procedure until all 15 have been arranged in order of merit. Give the whole series a final survey and rearrange the advertisements further if you wish.

Make out a table with narrow columns for recording Page No. and Rank No., a wide column for recording specific reasons, and a fourth narrow column for general reasons. Divide the table into wide rows. Fill in the rows of the first column with numbers 1 to 15. Record in the second column the page number of the advertisements judged to be of corresponding rank.

Write in the first row of the third column what you consider to be the specific *reasons* for No. 1's catching your attention so strongly. Do not be satisfied with any single reason, but state several reasons why you think the advertisement caught your attention at first glance, why it was an especially attractive one to you personally and would be so to most people in general. Record only what makes it catch (not hold) attention. Try to give two or more specific reasons for each of the other advertisements.

Try now to classify these specific reasons under more general heads. Wherever it is at all possible, indicate these general reasons in the fourth column by numbers which refer back to the list of 11 conditions given in the text above. Note which specific reasons seem impossible to classify under one or more general headings.

Part 2. *Value for Holding Interest.*—Jumble the order of the pages and lay out the advertisements again. Make out a table after the model of the table in Part 1. Take the first two advertisements you come to, look at and read each of the two advertisements so long as it

interests you in the *slightest degree.* Which one interested you more and held your interest more closely? Place this at the left in position No. 1 and place the other in position No. 2. Now take a third advertisement, look at and read it so long as it interests you in the slightest degree. If it interested you longer than the advertisement in position No. 1, place it to the left of No. 1. Otherwise compare it with No. 2 and place it between No. 1 and No. 2 or to the right of No. 2 according as to whether it interested you for a longer or shorter time than No. 2. In this way complete the ranking of the advertisements for holding interest and record ranks in table. If all were of some interest record the fact. If some were not, draw a double line across the table below the last recorded. Then force yourself to study the remaining advertisements until you have examined each quite fully. Judge and record the relative interest values where interest was forced.

Fill in the third and fourth columns with specific and with classified general reasons as in Part 1.

Report to the instructor the rank order for catching, and the rank order for holding attention of each advertisement. He will record your results in two tables which show for each advertisement the ranks reported by every member of the class. If the class is large and composed of both sexes, he will copy men's records in the first half and women's in the second half of the table. Copy these tables of class records for computation and for comparison with your own individual results.

Results and Discussion.

(1) Summarize your difficulties in assigning specific

reasons and in classifying specific reasons under general headings. Are any additions suggested to the list of general conditions listed in the introduction? If so, what are they?

(2) Compute the median and AD from each median for ranks assigned each of the advertisements in rating for tendency to catch attention. On what advertisement do judgments of the class agree best (most consistently)? Agree least well? Is there definitely closer agreement in rating advertisements of high, of low, or of average value? If so, try to explain.

(3) Answer the questions of (2) for the data referring to holding interest. Compare your answers with those of result (2).

(4) Correlate the ranks for catching attention assigned by yourself with the medians of the ranks assigned by the class for the same characteristics. State clearly and definitely what the coefficient tends to show. Obtain PE_ρ. Why is it rather large? What does the fact that it is large mean?

(5) Correlate the ranks for holding interest assigned by yourself with the medians of the ranks assigned by the class for the same characteristic. Compare the coefficient thus obtained with the coefficient obtained in result (4). What does this comparison tend to show?

(6) If the class data are of a sort and so arranged as to make it possible, obtain the median rank for attention value assigned to each advertisement by men. Make a similar computation for the ranks assigned by women and then correlate the two class rank orders so obtained. State clearly just what the coefficient tends to show.

(7) Mention some qualities in an advertisement which you think ought to be especially avoided.

Mention two examples of poor or bad advertisements and give psychological reasons why you consider them so.

In addition to magazine advertisements, mention at least four rather different ways of advertising. What particular psychological or other advantages do you think each one of the four ways possess?

(8) Try to mention a specific advertisement which you think appeals particularly to each one of the 11 native or acquired interests listed in the introduction.

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 365-381; J. F. Dashiell: *Fundamentals of Objective Psychology,* 1928, pp. 284-303; E. B. Titchener: *Textbook of Psychology,* 1911, pp. 265-302; A. T. Poffenberger: *Psychology in Advertising,* 1925, pp. 146-300, esp. pp. 146-173; D. Starch: *Principles of Advertising,* 1923, pp. 615-669.

CHAPTER XXXII

Experiment No. 26

THE AFFECTIVE VALUE OF COLORS

There has been an increasing use of color and color combinations in advertising. Evidence indicates that colored advertisements are much more effective than black and white. The effectiveness of color seems to lie partly in its contrast effect with black and white, and partly in its artistic and illustrative value. One important justification for the use of colors in advertisements is the pleasant feeling-tone which they arouse. Single colors as well as color combinations have affective value.

To clarify our discussion the meaning of certain technical terms in the field of color vision and certain facts of color experience are explained below. Hue is the term used to designate the quality of a color. Red, green, and blue are hues. Value or brightness of a color designates the degree of darkness or lightness, ranging from very dark through the spectral color to very light. The dark colors are called shades of a color and are produced by mixing black with a spectral hue. When white is mixed with a spectral hue a light color results and is termed a tint. For example, the addition of black to blue yields the shade navy blue; the addition of white to red, the tint pink. Chroma and saturation are identical in meaning, and are terms employed to signify the amount of color present. A fully saturated red is a strong, rich red in contrast to weaker saturations having

the same brightness. The weaker saturation is produced by mixing the saturated red with a gray whose brightness is the same.

Two colors which, when mixed in the proper proportions, cancel each other and produce a neutral gray are called complementary colors. The following are illustrations:

Color	*Complement*	*Color*	*Complement*
Red	Blue-green	Orange	Green-blue
Blue	Orange-yellow	Yellow	Violet-blue
Violet	Green-yellow	Green	Violet-red

Colors as they appear in the spectrum are standard colors.* Mixtures of two or more of the four psychological elementary colors; red, green, yellow, and blue, together with black and white will give all the variations of color that the eye is capable of seeing.

Experimentation has shown that affective value depends on the hue, saturation, and brightness of colors. The hue blue is usually the most preferred color. In general, saturated colors are preferred to weak colors, either tints or shades, and bright colors to dull ones. Color preference is influenced by both age and sex. Throughout all school grades blue is very constant in holding first place and red in second place. In the lower grades yellow stands high and green low, while in the upper grades and for adults green is high and yellow low.

The method of paired comparisons, which we shall employ in this experiment, measures the relative value and not the absolute value of the members composing a series in respect to some particular quality or characteristic. As the name of the method implies, its proce-

* It is advisable for the instructor to illustrate these facts of color experience by charts and color mixture.

dure consists of presenting the stimuli in pairs. It provides that all possible pairs shall be presented for judgment in respect to some single quality so that *each* stimulus shall be compared with every other member of the series at least once. After each pair is presented a comparative judgment is made. The judgment of each pair results, then, in one "vote" or preference for the favorite member, and each stimulus in the series has an equal chance to receive a large number of preferences. The total number of preferences received by each member of the series constitutes the final measure. These total numbers are usually expressed in tabular form.

If the two stimuli of the pairs are colors, they may be simultaneously presented. If they are tones they are given in immediate succession and the necessity for memory thus be reduced to a minimum. In the former case, judgment may be prejudiced by the fact that one stimulus is on the right or left, or one above and the other below (space errors); in the latter case by the fact that a given stimulus is presented first in a pair (time error). To cancel these errors each pair of the series of trials should therefore be presented twice, any given stimulus appearing once first (or on the right), and once second (or on the left) in each pair. The number of trials in the whole experiment will thereby be doubled, and the general result made more reliable.

Suppose there are three stimuli in the series—A, B, C. The space or time order of pairs and the number of trials necessary for such a series may be represented thus: A-B; B-C; C-A; A-C; C-B; B-A. A similar principle of arrangement would of course be followed in case a greater number of stimuli were involved. Note that the order of the last 3 pairs is a reversal of the order of the

first 3 pairs, with the additional provision that the succession within each pair is also reversed, thus compensating for practice as well as for constant error.

Any sort of stimuli whatever may be compared by this method, provided only that they are of such a nature that all possess some single quality on which reasonably consistent comparative judgments may be passed. We might, for example, thus compare a series of colors with respect to saturation, or with respect to pleasantness; a series of statements as to their truthfulness; a series of crimes as to their wickedness; a series of advertisements for their attention compelling power; etc.

The purpose of this experiment is: (1) to obtain color preferences for single colors and color combinations by the method of paired comparisons and to compute sex differences for these preferences; (2) to compare preferences for single colors to preferences for combinations; and (3) to note applications of these findings to advertising.

Apparatus.—Timer or stop watch; a series of six numbered cardboards, each containing one of the six following colored papers: R, G, B, Y, O, V; a series of fifteen numbered cardboards, each containing two colored papers, each of the above six colors is paired with every other color; an exposure frame (see Appendix).

Method.—The instructor serves as experimenter and the students as subjects. The method used in this experiment is known as paired comparisons. A pair (or 2 combinations) of colors are exposed side by side and the subject is to state which of the two he prefers. Two tables are prepared by each subject. The first is labeled "Single Colors" and has 15 blank spaces numbered 1 to 15. The second table is labeled "Color Combinations"

and has 105 blank spaces numbered 1 to 105. The number of comparisons is determined by $\frac{n(n-1)}{2}$ where n is the number of stimuli in the series.

Part 1. *Single Colors.*—Instruct subjects: "*Two colors which are numbered will be exposed side by side for a brief interval. Decide which of the two exposed colors you prefer and put the number of that color down in the appropriate place in your table.*"

The experimenter says "ready" before exposing each pair. Be sure that instructions are understood before proceeding with the experimental series. Expose the colors for about three seconds by lowering and raising the cardboard cover on the exposure frame. To avoid possible time and space errors use the procedure for exposing colors outlined in *Notes for Instructors.*

It will be noted that each color comes twice in succession; the first experiment compares colors (1) and (2); the second, colors (1) and (3); the third, colors (2) and (3); and so on. Vary the position of each color from experiment to experiment. Let color (1) be on the right in experiment 1, and on the left in experiment 2; color (2) on the right in experiment 2 and on the left in experiment 3. This means that the color which is left on the right side of the exposure frame after one exposure is shifted to the left side for the succeeding experiment. When the above series of experiments are completed, each color will have been compared with every other one.*

* If it is not convenient to use an exposure frame the instructor can expose stimuli by holding up the colors side by side. It is a little more difficult to do accurate timing by this technique but slight variation in exposure time does not seem to influence results in this experiment.

Part 2. *Pairs of Colors.*—Instruct subjects: "*The procedure here will be just the same as in Part 1 but 2 pairs of colors will be exposed instead of two single colors. As before, write down the number of the preferred combination in the appropriate place in your table.*"

The experimenter uses an outline of procedure for exposing color combinations similar to the one employed above with single colors. (See *Notes for Instructors.*) Using the 15 pairs of colors will yield 105 comparisons. Follow this outline of procedure and vary the position of each color from experiment to experiment as in Part 1.

In scoring, have each student count up the number of times each number (color) occurred in his responses: (A) for single colors, (B) for color combinations. Tabulate the results on the blackboard as follows, keeping the sexes separate:

SINGLE COLORS

Subject No.	COLORS FOR MEN					
	1	2	3	4	5	6

Subject No.	COLORS FOR WOMEN					
	1	2	3	4	5	6

Make a similar table for color combinations and head the columns 1 to 15. The color key for the numbers fol-

lows: 1 is red, 2 is blue, 3 is green, 4 is yellow, 5 is orange, 6 is violet, for the single colors. For the color combinations: 1 is B-R, 2 is G-R, 3 is R-Y, 4 is O-R, 5 is V-R, 6 is B-G, 7 is B-Y, 8 is O-B, 9 is V-B, 10 is G-Y, 11 is O-G, 12 is V-G, 13 is O-Y, 14 is V-Y, 15 is V-O. The data are to be copied by the students for their computations.

Results and Discussion.

(1) Compute the number of times each single color was preferred by the men; by the women. Giving the most frequently preferred color a rank of 1, rank the colors for each sex according to preference and tabulate results. What sex differences appear in color preference for single colors? Obtain the correlation between the ranking for men and that for women and the PE_ρ. Why is the PE of this coefficient of correlation high?

(2) A college class found that men preferred single colors in the following order: Blue (most), red, green, violet, orange, yellow. Women in the following order: Blue (most), red, green, yellow, orange, violet. Point out similarities and differences to show how the results of our class agree with those of a typical class.

(3) What is the difference in color preference for isolated colors and for colors in advertisements (Poffenberger, p. 446)?

(4) Answer each of the detailed questions in problem (1) for color-combinations.

(5) A college class obtained the following order of preferences for color combinations.

	Men	*Women*		*Men*	*Women*		*Men*	*Women*
B-R	2	6	B-G	1	4	O-G	7	9
G-R	6	10	B-Y	4	1	V-G	12	11
R-Y	8	7	O-B	3	5	O-Y	10	8
O-R	9	12	V-B	13	14	V-Y	14	3
V-R	15	15	G-Y	5	2	V-O	11	13

Correlate the results of our class with these typical results and state what the coefficient for the men tends to show. For the women. Would you expect to find a high correspondence between our class results and the typical results? Why?

List several factors which might affect one's prefence for colors.

(6) Geissler states that: "The greater the pleasantness of the individual constituents, the greater will be the pleasantness of the combination." How well do our results support this assertion? Cite evidence to support your statement.

(7) Plan a color scheme for advertising: (A) a perfume, (B) an automobile. Keep in mind not only the "appeal" of the color combinations themselves, but also their suitability for expressing qualities that may be associated with the object to be advertised. (For suggestions, see Poffenberger, pp. 455-458.)

(8) Cite other methods which might be used to study color preferences. Compare their merits with those of the paired comparisons method. What is meant by "space" errors? How were they controlled in this experiment?

Summary and Conclusions.

References

For color in advertising, see A. T. Poffenberger: *Psychology in Advertising*, 1925, pp. 428-463; D. Starch: *Principles of Advertising*, 1923, pp. 579-605; A. J. Brewster and H. H. Palmer: *Introduction to Advertising*, 1927, pp. 185-192.

For color vision, see R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 342-361; B. B. Breese: *Psychology*, 1921, pp. 178-186; A. I. Gates: *Elementary Psychology*, 1928 (Rev. Ed.), pp. 148-156.

CHAPTER XXXIII

Experiment No. 27

THE AFFECTIVE VALUE OF MUSICAL INTERVALS

Review the procedure of the order of merit method, the paired comparisons method, and graphic rating on an absolute scale (Experiments Nos. 23, 25, 26).

The measurement obtained by the order of merit *method* is a relative estimate, and not an absolute measure of the excellence or value of the stimuli in the series. The same is true of the paired comparisons method. In the former method no definite limit is set to the *time* which S may take in appreciating any given stimulus. He may, on the contrary, neglect almost entirely or fail to appreciate sufficiently certain members of the series through too brief notice, or he may over-appreciate other members through too extended attention to them. Also in the order of merit method there is no *controlled order of presentation,* and usually no attempt is made to counteract or to measure *errors due to spacial position and the like.*

On the other hand, in the method of paired comparisons, stimuli are presented for a certain definite interval, the spacial order of presentation is controlled to counteract possible space errors, the time order may be varied to control possible time errors, and the ABBA arrangement of the stimulus series may be employed to compensate for practice effects. Although the paired compari-

sons method seems to lend itself to better experimental controls than the order of merit method, practically identical results are obtained when the same series of stimuli are presented by one method and then the other. Barrett, using three types of material, found the average correlation between the two methods to be .987. Choice between the two methods, therefore, should be made on the grounds of adaptability to a certain situation or convenience, such as economy of time, etc. The paired comparisons method appears to be best adapted to class experiments and the order of merit method seems best for individual experimentation where economy of time, avoidance of fatigue and monotony are considerations. When the series of stimuli contains 10 or more objects the paired comparisons method involves a large number of comparisons. The number of comparisons increases rapidly with increasing length of the series. This prohibits the use of this method in the laboratory for practical purposes whenever the number of stimuli in the series is much greater than 15. With auditory stimuli, of course, the order of merit method cannot be employed.

With the graphic rating procedure, when employed to rate personality, achievement and intellectual traits,* there are certain important sources of error. When used for measuring the value of certain stimuli, as the affective value of musical intervals, the graphic rating scale appears to be less subject to errors of measurement than when used for personality traits. This is probably due to the nature of the situation and the kind of stimuli employed. With musical intervals the stimuli may be repeated in such a manner as easily to multiply the number of measurements without the rater being influ-

* See Burtt reference below.

enced by previous ratings. Rating one's degree of preference for a musical interval is also apparently more objective than rating a personality trait. Rating the affective value of 12 musical intervals is somewhat analogous to rating 12 people on a single trait.

The purpose of this experiment is (1) to obtain and compare the individual and group preferences for the intervals of the diatonic musical scale and relate these preferences to size of intervals and degree of consonance; (2) to determine the relationship between two methods for ascertaining preferences: (a) the paired comparisons method, and (b) an absolute rating scale method.

Apparatus.—The Ellis Harmonical * with untempered tuning, or any piano or organ tuned in ordinary fashion (equal temperament); check sheets for scoring (see *Notes for Instructors*).

Method.—The instructor will act as E and all members of the class will serve as subjects.

Part 1. *Paired Comparisons.*—The formula employed to determine the number of comparisons in the paired comparisons method depends on the need for controlling possible time errors. Where the tendency for time errors to occur is negligible the formula is $\frac{n(n-1)}{2}$; where it is desirable to control possible time errors, it is $n(n-1)$, in which $n =$ the number of stimuli. Each member of the class first makes out a two-column table of $n(n-1)$

* The Ellis Harmonical is a species of harmonium or reed organ designed and tuned to give a perfect scale, and which produces 24 overtones with the tone of 66 vibrations and 16 overtones with the tone of 132 vibrations, etc. Because the air is compressed in the bellows and conveyed thence to the reeds, it has a harsher tone than the ordinary harmonium which uses suction.

= 132 rows * in which the trial numbers and the judgments are to be recorded.

The instructor (or assistant) sounds the chords in the order indicated in a check sheet (see *Notes for Instructors*). Each chord should be sounded for about two seconds with a two-second interval between, and a five-second interval between successive pairs. It is highly important to strike the dyads (chords) with uniform intensity if one is to obtain consistent results. The number of the trial should be announced in advance of each comparison.

The subjects are instructed to maintain a steady passive attitude throughout the series of trials. After each pair of intervals they judge: "Which one of the intervals (chords), the first or the second, is *more pleasant?*" and record this judgment as "1" or "2" in the appropriate place in the table. Judgments of "=" should never be given.

Judgments should be naïve and *"immediate,"* that is to say, not reflective or "thought out," but based simply upon the immediate effect produced by the fusions.

After completion of the 132 trials of the experimental series, compare the tables of judgments with a checking table, furnished by the instructor, which shows the actual order of intervals used. Compare this list with your response list, underlining the intervals which seemed more pleasant to you (thus: *octave*—major 2). From the underlined list sum up the number of preferences (votes) for each interval and express the results in a third table: octave has (—?) preferences; fifth has (—?) preferences; etc. When each student has listed his pref-

* Since there is apparently only a slight tendency toward time errors in this experiment the instructor may employ only the first 66 comparisons of the series and still obtain highly reliable results. If the whole series is used, give the students a rest period of about 10 minutes after the 66th trial.

erences, he should tabulate his contribution to the class data on the blackboard or on a sheet prepared by the instructor. Construct this table with 12 columns, each column headed with the name of an interval. When the data are all tabulated the instructor or assistant is to sum up each column and the students are to copy these sums, properly labeled, for use in their computations.

Part 2. *Rating on an Absolute Scale.*—Each member of the class is furnished with a rating blank which provides for 36 ratings (see *Notes for Instructors*).

Each interval occurs 3 times in the series. The instructor sounds the 36 chords in the order indicated in a check sheet. Each chord should be sounded for about two seconds with a 5-second interval between successive chords. Strike the chords with uniform intensity. The number of the trial should be announced in advance of each stimulus. Instructions on the rating blank read: "*For each interval, a graphic rating line is given. If the effect of the musical interval seems very pleasing, put an X at the extreme right of the line. If it seems very displeasing put an X at the extreme left of the line. If the degree of pleasantness or unpleasantness falls between these two extremes, place the X at the point on the line that best expresses your liking or dislike. The subject need not mark directly over a descriptive adjective but may mark anywhere along the line. Rate each musical interval in the sequence in a similar manner and on its appropriate line. The descriptive phrases for guidance in your rating are given only under the first and last line on each page of the blank. Make your rating each time just as though the phrases were repeated for each line.*" Give two or three samples to illustrate the method.

After completing the 36 trials of the experimental series, compare ratings with a checking table, to be furnished by the instructor, which shows the actual order of intervals used. On each rating line measure the distance of the X from the left end in millimeters or sixteenths of an inch (ask instructor which to use). Copy these distances opposite the appropriate intervals on the checking sheet and then compute the average (of three) distance for each chord: octave has (—?) millimeters, etc. The longest average distance designates the most preferred interval, the shortest the least preferred. In a table similar to that used in Part I tabulate class data, add up the figures in each column and have students copy these sums for their computations.

Results and Discussion.

Part 1.

(1) Assign rank orders for pleasantness to your own and to the class results. Correlate the ranks, obtain PE_ρ, and interpret the obtained coefficient. Point out any marked difference between your ranking and that for the class data and try to explain the difference.

(2) The size of the intervals, in terms of vibration difference between the components of each chord, varies from large to small as follows: octave, major seventh, minor seventh, major sixth, minor sixth, fifth, tritone, fourth, major third, minor third, major second, minor second. Giving the largest interval a rank of 1, how does the rank order for size of interval correlate with the class rank for pleasantness? Compute PE_ρ. What does the correlation mean?

(3) Centuries ago, the octave always had the highest rank for pleasantness. How would you account for this fact, and how account for its subsequent drop?

(4) The rank order of the intervals for decreasing degree of consonance (completeness of fusion) is: 1. octave (greatest consonance), 2. fifth, 3. fourth, 4. major sixth, 5. minor sixth, 6. major third, 7. minor third, 8. major seventh, 9. major second, 10. tritone, 11. minor seventh, 12. minor second. Correlate this ranking with the class rank order for pleasantness and interpret the coefficient.

(5) Try to describe the feeling of pleasantness. That is, try to tell how you feel; what you do; or how you behave when a fusion pleases you (or when some other type of stimulus pleases you). Do you feel the pleasantness in any particular part of the body? etc. Similarly for unpleasantness. Which theory of feeling does the answer you have just given support (see Gates, pp. 176-178)? Why?

(6) Compare your class rank order for pleasantness with the following ranks obtained from a representative group of college students: 1. major third (most pleasant); 2. minor third; 3. minor sixth; 4. major sixth; 5. fourth; 6. tritone; 7. fifth; 8. octave; 9. major second; 10. minor seventh; 11. major seventh; 12. minor second. Make out a list of differences between ranks and state which intervals are judged with greatest constancy from group to group; which with least constancy. Do those which are most and least liked show greater constancy than those near the middle of the list? Is there any reason for this? What effect would the

inconsistencies in ranks have on the correlation between the two series?

Part 2.

(7) Correlate class results from Part 2 with those from Part 1. What relationship is discovered? Which method of determining preferences do you consider the better, and why? Which method did you like the best and why?

(8) What procedure would you follow to compute the reliability of the test in Part 1? In Part 2?

Summary and Conclusions.

References

For discussions of feelings, see R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 280-285; A. I. Gates: *Elementary Psychology* (Rev. Ed.), 1928, pp. 171-180; E. B. Titchener: *Text Book of Psychology*, 1911, pp. 225-264; *Experimental Psychology*, Vol. 1, Part 2, 1901, pp. 149-158; B. B. Breese: *Psychology*, 1921, pp. 356-376; W. B. Pillsbury: *Fundamentals of Psychology*, 1922 (Rev. Ed.), pp. 458-479.

For methods, see M. Barrett: "A Comparison of the Order of Merit Method and the Method of Paired Comparisons," *Psychol. Rev.*, 1914, Vol. 21, pp. 278-294; P. R. Farnsworth and C. F. Voegelin: "Dyad Differences at Different Intensities," *Jour. Appl. Psychol.*, 1928, Vol. 12, pp. 148-151; H. E. Burtt: *Principles of Employment Psychology*, 1926, pp. 317-370; W. V. D. Bingham and M. Freyd: *Procedures in Employment Psychology*, 1926, pp. 122-142.

CHAPTER XXXIV

Experiment No. 28

THE "FREE" ASSOCIATION OF IDEAS

The *general law of association* states that if two ideas have at any time occurred together or in immediate succession the later recurrence of the one *tends* also to bring about the recurrence of the other. The basis of this tendency is to be found in the nervous system, upon whose action all psychological prediction and explanation ultimately depend.

In terms of nervous functions, the law may be restated as follows: If two groups of neurons (corresponding to two ideas or perceptions) have been excited together, the re-excitation of one group tends to be followed by re-excitation of the other. This is the general law of habit, applied specifically to the case of ideas.

Since any given idea (or at least any element of an idea) has occurred in the past in connection with a great variety of others, all of these others, according to the law, are *likely* to recur. The general law, therefore, has little value for prediction, and even approximate prediction will evidently require a knowledge of *secondary* laws of a much more specific character. Such secondary laws are, for example, (1) that the most *frequently* associated idea, (2) that the most *vividly* experienced idea, and (3) that the most *recently* associated idea, has the greatest tendency to reappear. Some sixty other secondary laws of association of greater or lesser impor-

tance and degree of specificity may be discovered in the psychological literature. Some of these secondary laws have already been illustrated in our experiments on memory.

But association through recency, frequency, and vividness and the other factors mentioned does not in everyday life *completely* determine the sequence of ideas. It is evident to casual observation that in normal thinking, for example, we *"set"* ourselves in a particular direction, we confine our ideas to a certain *topic,* set out to solve a given *problem* or to accomplish a definite *purpose.* In such cases, the purpose or problem, if we are at all concentrated upon it, brings it about that ideas which are *relevant,* shall appear, and that irrelevant ideas, no matter how strongly associated with our present one, rarely arise.

It should be noted also that the "set" or purpose need not always be one of which we are conscious. Once we get well into a problem we may become absorbed in it and forget why we are trying to solve it, but the nervous mechanism of the set, once established, may still persist and "control" the selection of ideas. The point here, then, is not that ideas appear without involving association, but simply that the strongest associations are not necessarily effective when they fail to fit into our "set," or fail to meet the demands of our purposes. Otherwise stated—the sequence of ideas is determined in part by our associations (habits), but in part also by our "set."

The associations in directed thinking are sometimes called "constrained" or "directed" associations as opposed to the "free" or relatively undirected associations which are determined primarily by habit.

If a person is told to think out and to report rapidly as many different words as possible (not in a sentence)—"grass, table, yellow, go"—these verbal ideas will to a large extent be free associations. Even in such a case, however, we often find that groups of words belong to a number of varied topics. To avoid topical associations in so far as possible, the standard experiment in free association consists in calling out to a subject in order a large number of very different stimulus-words, and in every case requiring S to respond as quickly as possible by giving the *first* response that the stimulus-word suggests. It is such an experiment that we shall here perform.

Since the training, interests and experiences of everyday life are in many respects alike for normal adults, it is to be expected that in this free association experiment many identical response words will occur to different individuals. Otherwise expressed, normal individuals may be expected to show a certain *community* of association responses.

P. E. Schellenberg secured from each of 925 college freshmen, men and women at the University of Minnesota, a series of association responses to 100 stimulus words.* Complete frequency tables for the responses to

* These are the 100 words employed by Grace Kent and A. J. Rosanoff in securing from each of 1000 normal subjects a series of association responses. For each stimulus word in the series they compiled a frequency table to show the number of times in 1000 that each response occurred. Ordinarily the frequencies in this table have been used as norms with which the association responses of subjects are compared. Certain distinct limitations detract from the usefulness of the Kent-Rosanoff frequency table when working with contemporary college students: (1) the association responses were obtained from groups with many degrees of intelligence and education. (2) The subjects ranged in age from 8 to 80 years. (3) The frequency table was compiled in 1910. These limitations are emphasized by the fact that Schellen-

each stimulus word were made by recording all the response words and by counting the number of students who gave each response word. The frequency tables which we shall use contain only the three words which were given most often in response to each stimulus word; the numbers indicate the actual number of students who replied with each response. By comparison of our subjects' responses with those in this frequency table we may determine mathematically the likeness of our S's associations to those of a group of university freshmen. Experimental evidence also shows that individuality of response indicates high intelligence (i.e., there is a negative correlation between intelligence test scores and community of response scores). Greater individuality of response of course, is indicated by lower frequency scores. The reason for this is that the higher the frequency score the more people there were giving that response. In addition to the frequency of response the time of association reaction may also be significant. It has been found, for example, that words associated with unpleasant or exciting emotional experiences ("com-

berg's frequency list, when compared with the Kent-Rosanoff table, shows a total of 91 new responses added to the three most frequent responses, and a total of 83 changes from one position to another (first to second, etc.), in the list of three most frequent responses. In the shorter form of Schellenberg's frequency table which we shall use in this book, are given merely the three most frequent responses to each stimulus word. This condensed form was made because J. Clousing had previously shown that scoring on the basis of the three most frequent responses in the Kent-Rosanoff Frequency Table yielded results which correlated .95 to .99 with scores derived from the whole frequency table. Thus scoring from this abbreviated list yields results which are practically identical with those from the complete table. Use of a short table avoids the very laborious task involved in scoring from the complete Kent-Rosanoff frequencies. Moreover, since Schellenberg's frequency table was standardized on a college population it should be more valid for use where college students are subjects.

plexes'') of the individual frequently show lengthened reaction-times. Lists of stimulus words, to be used under special conditions and with special instructions for free association have been designed to assist in diagnosing such complexes. The bases of judgment in such cases are not only reaction times, but also certain other ''complex'' indicators, some of which are discussed in the succeeding experiment.

The purpose of this experiment is to demonstrate and measure roughly the community of response in the free written association reaction of subjects by comparing their responses with those of 925 college freshmen.

Apparatus.—Schellenberg Frequency Table,* stop-watch or timer, list of 100 stimulus words. (See *Notes for Instructors.*)

Method.—The instructor serves as experimenter and the class as subjects. The 3-page mimeographed folders containing the stimulus words and the directions are distributed to the students after they are warned not to turn any pages without instructions to do so. The subjects are then directed to put name, date and sex in the appropriate places on the first page. The instructor reads aloud the instructions printed on the first page: ''On the following two pages you will find lists of words. *Write after each word the very first word* (*not a phrase or sentence*) *that comes to your mind.* Be sure to write legibly. Do this as rapidly as possible because this is a *speed test.* As soon as you finish writing the last word turn back to this page and in the space following 'time' copy the figures that appear on the blackboard at that moment. This will be a measure of the time you re-

* The Kent-Rosanoff frequency table may be used if the instructor desires. In which case compute medians rather than averages in results.

quired to complete the task. Wait until the signal is given before turning this page."

"Are there any questions?" (Illustrate a stimulus word and response on blackboard.) "Ready, begin." The timer is started and, after 2 minutes (2′ 00″), the time is recorded on the blackboard every 15 seconds. The preceding figure should be erased each time a new one is recorded.*

Rough notes should be written out for answering question (6) while the experiment is fresh in S's memory.

After the 100 responses have been obtained and the total time recorded S obtains a copy of Schellenberg's frequency table from the instructor. By means of this table he looks up the frequency of every response and enters this in the appropriate column and row of his record. If S's response word does not appear in the tables at all, it is assigned a frequency of zero. The student should note that zero frequency on the abbreviated Schellenberg list does not necessarily mean that none of the 925 college students gave this word. It simply means that it does not appear among the three most frequent responses. The word may be fourth highest and still not be on the list.

S should report to the instructor the total time in minutes and seconds taken to write the 100 association responses, the average of the response frequencies of the first 50, the second 50, and the whole 100 words. He also reports the number of zero responses for the entire list and copies from the blackboard similar data obtained from the class.

* Code letters may be used instead of writing the actual time on the blackboard.

Results and Discussion.

(1) What is the range of average response frequencies of the 100 responses for the class? Would you expect a comparatively large or small range for an ordinary college class of 30 or 40 students? Why?

(2) What is the average of the response frequencies of your first 50 responses? Of second 50 responses? How much difference is there between these two averages? In your own case, therefore, would the responses to 50 words seem to yield a reasonably reliable measure of community of ideas, or need responses for at least 100 words be taken to increase the reliability of the measure?

(3) Using class data for the whole 100 words correlate speed of response with individuality of response. (The subject having the shortest response time, i.e., the speediest S, receives a rank of 1; the S with the lowest average frequency shows greatest individuality of response and is assigned a rank of 1.) State clearly just what the obtained coefficient tends to show concerning correspondence between speed of response and individuality of response.

(4) How many of S's (your own) responses were found to have zero frequencies? What might a large number of such responses indicate? Explain fully.

Could one use the number of zero responses as a measure of community of response? Correlate zero frequencies with average frequencies, giving S with the least number of zero responses a rank of 1 and S with the highest average response a rank of 1. Why this arrangement? Explain clearly just what the obtained coefficient tends to show.

(5) One way of computing a reliability coefficient for a

measure is to correlate scores in the first half with scores in the last half of the test. Correlate the averages of the response frequencies (class data) for the first 50 words with the averages for the second 50. With what degree of assurance can one predict scores on the last half of the test from a knowledge of scores on the first half? Would you advise using merely the first 50 words for the test? Why?

(6) Describe as well as you can what takes place consciously between stimulus and response. Do you, for example, commonly find imagery of any given sort appearing before you respond? Do you have "blockings," etc.?

Select a few particular responses which you think are due primarily (a) to recency of association, (b) to frequency of association. Give reasons for your selections.

(7) Would you expect the responses of an insane person to show a greater or less average frequency (community) of responses than that of an individual in this class? (If you need help, see Kent and Rosanoff, p. 317-390.)

(8) Describe briefly three other methods of measuring or determining character and personality traits. (See Gates, pp. 533-543.)

Summary and Conclusions.

References

R. S. Woodworth: *Psychology,* 1929 (Rev. Ed.), pp. 427-431; A. I. Gates: *Elementary Psychology,* 1928 (Rev. Ed.), pp. 533-543; J. F. Dashiell: *Fundamentals of Ob-*

jective Psychology, 1928, pp. 221-223, 278, 545-547; G. H. Kent and A. J. Rosanoff: "A Study of Association in Insanity," *American Journal of Insanity,* Vol. 67, 1910, pp. 37-96, 317-390; P. E. Schellenberg: *A Free Association Group Test for College Students,* a Ph.D. thesis at the University of Minnesota; J. Clousing: *The Relation Between Personality Ratings and Free Association Responses,* M.A. thesis, 1927, University of Minnesota.

CHAPTER XXXV

Experiment No. 29

ASSOCIATIVE RESPONSES AS DIAGNOSTIC

In the experiment on "free" association we saw that the character of such associative verbal responses depends to a large extent upon the past experiences and upon the present interests of the subject. We found that because of the similarity of the past experiences in normal subjects community ("commonplaceness") of responses is the rule.

We noted, however, that frequency of association in past experiences is not the only determinant or conditioning factor of associative reproduction. If a subject has a particular and more or less permanent line of interest, or if he has a temporary "set" due to recent instructions or suggestions, the responses become "controlled" or constrained associations. The delusions, obsessions, and fixed ideas ("complexes") of the insane and the "hobbies" of normal persons are thus likely to determine the character of certain responses. The list of stimulus-words prepared by Jung contains words especially likely to betray such complexes (see B. Hart: *The Psychology of Insanity,* 1914, esp. pp. 58-76). And, furthermore, the influence of vivid thought or experiences (especially if they are recent or emotionally toned) may overweigh that of frequency, even when the responses thus determined are rather loosely associated with the stimulus-words and would not under ordinary conditions provoke such responses.

Instead of using Jung's list, and trusting to chance that some of the words in it may refer to a complex in our subjects, we shall here endeavor experimentally to establish a complex in our subject and thereafter try to see if we can discover evidence of it through a study of his responses to a series of stimuli especially suited to betray that particular complex.

We may judge concerning the presence or absence of a particular complex by noting whether or not (a) a relevant stimulus-word brings forth a response which would not naturally be given by a person who had not recently had the emotional experience in question; (b) a relevant word fails to bring out such a response as would naturally occur to one who had had the experience; (c) significant responses are given with unusual frequency, perhaps even with irrelevant stimuli; (d) there is apparent a "forced" avoidance of relevant associations.

Complexes betray themselves not only through this "unnatural," "unexpected" or "individual" *quality* in the responses, but also by a variety of other diagnostic signs; (e) unusually long reaction times for responses to "critical" words; (f) unusually variable reaction times; (g) "far-fetched" associations and rhyming associations; (h) repeating as a response a previous stimulus or a previous response; (i) rapid forgetting of responses to critical stimulus-words.

Part 1. *Diagnostic Association as an Individual Experiment.*—The purpose of this experiment is to determine how well one can judge concerning the presence or absence of a particular "complex" by an analysis of various diagnostic signs.

Apparatus.—Stop-watch or timer, stimulus-list, boxes

with unknown contents hidden in a dark room. (See *Notes for Instructors* for details.)

Method.—One-half of the class are to serve as subjects and the other half as experimenters. The subjects are conducted (one at a time) into the dark room by the instructor. They are there told to select and to open one (and only one) of the two covered boxes, and to inspect very carefully and completely the contents of the box. They then report back to their partners for examination. *S is particularly cautioned that under no circumstances is he to admit to anyone which box he opened until he is given permission by the instructor to do so.*

When S returns have him write a description, stating in detail all that he can remember of what he saw in the box he opened. E is not to see this description until after the whole experiment is completed.

While S is in the dark room or writing this description, make out a 60-row tabular form with columns for recording: stimulus; response; time in seconds.

Seat S at a table distant enough from other pairs so that conversation cannot be overheard. Read aloud to him the exact words of the following instructions: "*I am trying to discover through your free associations what was in the box you opened in the dark room. Close your eyes and keep them closed. I shall give you one at a time in fairly rapid succession a series of stimulus-words. You are to respond as quickly as possible to each word, giving a free association—that is to say, the first single word that the stimulus-word suggests to you —provided that you do not think that first word would be a 'give away.' If you do fear this, you may give a word other than the first one suggested, but totally ir-*

relevant responses and any delay in responding may also help me to judge correctly.

"Remember, I shall judge both by the character of the response you make and by the length of time that it takes you to reply. Your task therefore is to reply quickly and at the same time to try not to betray what was contained in the box you opened."

If S is not sure that he understands perfectly E may read the instructions a second time, *but he is not to explain further in any manner. Strictly avoid any comment on S's replies.* Do not allow S to see that you have any suspicions as to which box was opened, until after the questions below have all been answered.

Take particular pains to speak the stimulus-words distinctly, but never repeat a word if S fails to hear. Before calling out each stimulus-word say "Ready." Then give the word and take the time of each of S's responses with a stop-watch, starting the watch as you call out the stimulus, and stopping it as quickly as S begins to respond. ("I did not hear" is a response and should be recorded as such and the time recorded.) Work fairly rapidly, and waste no time between responses except for the purpose of recording.

After the series has been completed get special instructions for a second trial and record results for it as you did for trial No. 1.

DO NOT YET DISCUSS HIS RESPONSES, YOUR SUSPICIONS, OR THE EXPERIMENT IN GENERAL WITH S!!

Results and Discussion.

Find out from the instructor what each box actually contained and mark the critical or significant words for

Box No. 1 by check marks before the numbers designated by the instructor. Mark the critical words for Box No. 2 by double checks. The remaining words are *non-critical.*

(1) Judging from the *nature of the response words* alone in trial No. 1, do you think S opened Box No. 1 or Box No. 2? Justify your judgment in detail. Ask yourself which responses to critical words appear to be perfectly "commonplace" and which appear not to be so, but rather unusual and determined by special recent experience.

(2) What is the *longest single reaction time to a critical* word in trial No. 1? Judging from this alone, which box do you think S opened? Why?

(3) What is the *average time* of response for the *non-critical* words in trial No. 1? For the Box No. 1 critical words? For the Box No. 2 critical words? Judging from a comparison of these three averages alone, which box do you think S opened? Why?

(4) What is the m.v. of the times of the *non-critical* words taken by themselves (trial No. 1)? of the *critical* words referring to Box No. 1? of the *critical* words referring to Box No. 2? Judging by comparisons of these alone, which box is indicated?

(5) Compare the responses of trial No. 1 with those of trial No. 2 for non-critical words. What percentage are identical? What percentage are identical in case of the Box No. 1 criticals? In case of the Box No. 2 criticals? Which box is indicated by a comparison of these percentages?

(6) What is the average time of the non-critical responses in trial No. 2 of the Box No. 1 criticals? of the Box No. 2 criticals? Judging solely from

the time of criticals in comparison with that of non-criticals in trial No. 2 which box do you think S opened?

(7) Did S exhibit diagnostic sign (h) or (i)? If so, which box is indicated?

(8) How many of the seven above answers indicate Box No. 1? Box No. 2? How many are ambiguous or indicate nothing? Judging on the basis of all these signs taken together, which box do you think S opened?

(9) Which box did S actually open? After you have made your decision get S to tell you which box he opened. Which of the diagnostic signs examined in results (1)-(7) gave a correct, an indeterminate or an incorrect indication? Which of the seven signs appears to be most trustworthy? Which appears to be next best, etc.?

(10) What sources of error and uncertainty and what other difficulties do you see in applying this method for determining guilt in criminal cases?

Summary and Conclusions.

Part 2. *Diagnostic Association as a Class Experiment.*—The purpose of this experiment is to determine how well one can judge concerning the presence or absence of a particular "complex" by an analysis of various diagnostic signs.

Apparatus.—Stop-watch or timer, stimulus list, unknown materials in a separate room. (See *Notes for Instructors* for details.)

Method.—The instructor serves as experimenter and two members of the class as subjects. The two subjects, selected by the experimenter, are conducted out of the

room by an assistant. There each one is given an envelope, directed to open it and to follow the enclosed instructions. After the injunctions have been carried out the subjects are brought back to the laboratory, one at a time, for examination. The subjects are particularly cautioned that under no circumstances are they to admit to anyone what they did or saw until they are given permission to do so by the instructor.

The rest of the method is the same as in Part 1 (designed for individuals working in pairs), except that in the class experiment the instructor is E and examines the two subjects. During the examination the subject is seated at the front of the room. His chair is so placed that although he is facing partly away from the class, still the instructor and students can see his profile. S is told to close his eyes and is cautioned to give the response in a loud, clear voice so that the class will be able to hear.

The experimenter instructs S: "*I am trying to discover through your free associations what you did while you were out of the room. Close your eyes and keep them closed. I shall give you one at a time in fairly rapid succession, a series of stimulus words. You are to respond as quickly as possible to each word, giving a free association—that is to say, the first single word that the stimulus word suggests to you—provided that you do not think that first word would be a 'give away.' If you do fear this, you may give a word other than the first one suggested, but totally irrelevant responses and any delay in responding may also help me to judge correctly.*

"*Remember, I shall judge both by the character of the response you make and by the length of time it takes*

you to reply. Your task therefore is to reply quickly and at the same time to try not to betray what you did while out of the room."

Each student is given two copies of the stimulus words; one for each S. In the blank spaces opposite the stimulus words all students and the instructor record the responses as given by S. The instructor also records reaction times. When the experiment is completed E dictates the reaction times for the class to copy. Responses recorded by students may be checked by reference to E's list.

When the first S has been given one trial (series of responses) he is conducted outside and the other S brought in. The same procedure is carried out with the second S, and then the first S is returned for his second trial.

This subject then leaves the room while the other S is given his last trial. Both subjects should be absent from the room while E explains the experimental situations to the class.

Results and Discussion.

After the stimulus words have been presented twice to each subject the instructor describes to the class the situation which was participated in by one of the subjects while the other S waited in another room. Critical words refer to the "active" situation. These critical or significant words in the stimulus list are marked by check marks before the numbers designated by the instructor. The remaining words are non-critical. The instructor is to survey some of the results and point out their diagnostic significance. The students then work out problems given in Part I. Instead of Box No. 1 and

Box No. 2, the students have S No. 1 and S No. 2.* One S participated and one did not participate in a certain situation. In results (8) the students are to state which S they believe to be "guilty."

Summary and Conclusions.

References

E. B. Titchener: *Text Book of Psychology,* 1911, p. 446; H. Münsterberg: *On the Witness Stand,* 1908, pp. 73-110; H. M. Leach and M. F. Washburn: *American Journal of Psychology,* Vol. 21, 1910, pp. 162-167; R. M. Yerkes and C. S. Berry: *American Journal of Psychology,* Vol. 20, 1909, pp. 22-37; H. R. Crosland: *The Psychological Methods of Word-Association and Reaction-Time as Tests of Deception,* University of Oregon Publication, Psychol. Series, Vol. 1, No. 1, 1929.

For a popular account of the use of diagnostic association, see "A Boy Trapper Who Crossed Canada for Revenge," *Literary Digest,* February, 1923, Vol. 76, pp. 46-48; and "The Boy Who Didn't Cross Canada for Revenge," *Literary Digest,* March, 1923, Vol. 76, pp. 38-42.

* In result (3) the average times are to be obtained for subject 1, non-critical words; for subject 1, critical words; subject 2, non-critical words; subject 2, critical words. For result (4) the AD's for each of the conditions in (3) are to be obtained. In result (5) the per cent of identical responses in trial 1 and trial 2 for non-critical words for subject 1 is compared with the per cent for subject 2. The same is done for critical words. Result (3) is repeated for trial 2 in result (6).

CHAPTER XXXVI

Experiment No. 30

JUDGMENT OF INTELLIGENCE FROM PHOTOGRAPHS

In ancient times it was believed that the physical and mental characteristics as well as the fate of an individual could be judged correctly by an astrologer from a "reading of the stars under which the individual was born." That such judgments are believed in or at least are desired even today is evident from the fact that advertisements of horoscopes persist in many magazines. In later times the exponents of physiognomy, of palmistry, phrenology, graphology and the like made similar claims for judgments of character based upon signs found in the body and face, the hand, the conformation of the skull, the handwriting, and in other objective characteristics and products. At present "character analysis" by a modified and combined physiognomical and phrenological system has a greater vogue, and pays larger returns to its advertisers.

A general criticism of such systems is easy. We know that false assumptions regarding the physical and the neurological basis of character underlie many of them. Examination shows that wide disagreement exists among exponents of the systems as to the choice and the significance of the signs considered to be diagnostic. The systems are based for the most part upon unsystematic observation and far-fetched analogies. The readings given are usually couched in vague and high-sounding

words which tend to impress the ignorant and flatter the common man, and the statements made are often so ambiguous and so general that they will apply with equal truth to a multitude of persons besides the particular one judged. Truly quantitative estimates are seldom ventured and when mere presence or absence of an opposable quality is asserted, it is rarely considered that correct assertion is likely in 50 per cent of the cases by chance alone. The desire to believe in and to use such a system for practical purposes favors unwarranted belief, and this belief is fostered by failure to observe negative instances, failure to note the possibility that judgments are often unconsciously based upon secondary criteria (such as previous external knowledge, the subject's behavior and speech, etc.), and by the general readiness of the inexpert and the uninformed to accept suggestions and not to be aware of the dozens of logical and scientific fallacies of belief which may enter into persuasive argument.

But such criticism is not conclusive evidence, and the universality and persistence with which in daily life we judge mental characteristics by facial expression may well make us inquire whether scientific test of such systems is not possible and desirable. May there not be "something" in such systems in spite of their evident inconsistencies and inaccuracies? And if something, how much? We have already indicated in the discussion of correlation that scientific test is in theory easy and in practice difficult. We must first agree upon definite and stable meanings for the terms we use; we must show that a real *consensus* of opinion exists among acquaintances, employers, teachers, and friends of the particular individuals later to be judged by the system

under investigation; we must insist upon quantitative or rank order estimates on large groups; and we must make sure that secondary criteria are excluded as a basis for estimates by the system, etc. No one of these prerequisites is easily obtainable. The best demonstration that we can give in the laboratory is to call for judgments based upon photographs of unknown individuals whose mental qualities have already been measured in as reliable a way as exists at present,—namely by a standardized scale of tests of intelligence. By correlations between the rank orders assigned by different individuals we can see how well judges agree, and by correlations between judges' estimates and the results of objective tests we can conclude something as to the reliability of facial expression as an indicator of what the tests measure.

The purpose of this experiment is to determine how accurately intelligence can be judged from photographs by the order of merit method.

Apparatus.—Photographs of children, reproduced in the Appendix of this book. Scissors.

Method.—Cut out the photographs in such a manner that the letter designating each photograph will remain attached to the picture. Then spread the photographs out on a table in serial order according to the alphabet with picture A at the left. Consider the first two pictures at the left and neglect the others for the time being. Which one of the two children, in your judgment, looks the brighter, i.e., the more intelligent for that age? Put that photo at the left in position No. 1, and the other in position No. 2 to the right of No. 1.

Now take any other photograph, compare it with the picture of the child in position No. 1. If the child in

the new picture looks brighter than No. 1, place the photograph to the left of No. 1. If, however, this new child looks less intelligent than the child in position No. 1, compare the picture with the child in position No. 2. If the child appears to be brighter than the one in position No. 2, place the picture between No. 1 and No. 2. Otherwise place it in position No. 3 at the right of No. 2. Continue this procedure until all the photographs have been arranged in order of merit with the child you consider to be the brightest at the left and the one you consider least intelligent at the right. Give the whole series a final survey and rearrange the photographs if you wish.

Record the letters corresponding to the photographs opposite ranks 1 (brightest looking), 2 (next brightest looking), 3 (third brightest looking), etc., and report your rank order to the instructor for recording on the blackboard. Copy the class judgments there recorded.

When through with the photographs place them in an envelope and turn it in to the instructor. The pictures will be kept in the laboratory for use in later years with second-hand books.

Results and Discussion.

(1) Describe your difficulties in judging. Was the judging unexpectedly difficult? What facial characteristics did you rely upon in judging?

Do you think judgments based upon sight of the actual children would be more accurate? Why?

(2) Which three photographs do you feel most assurance about? Which three photographs do you feel least assurance about? Get from the instructor a list of the I.Q's. Did you rank the photographs

about which you felt high assurance more accurately than the others?

(3) Correlate your rank order with the order obtained by objective tests (tests yielding I.Q's). Report your correlation to the instructor.

Copy from the blackboard the correlations obtained by members of the class. What is the median of these correlation coefficients?

From the data on ranks given by the class obtain for each child the median of the ranks assigned. Correlate this "class rank order" with the order obtained by the objective test (I.Q. rank order). What does this comparison show?

Discuss results obtained in this question and state whatever generalization appears warranted by the data.

(4) Correlate your ranking with that of the class (class-rank order obtained in result 3). Do you agree better with your classmates than you do with the results of the objective test? Try to explain. Report this correlation to the instructor for recording on the blackboard.

From the blackboard results find the median of correlations of each person with total class results. Compare with median correlation between judge's rank and I.Q. rank and explain.

Summary and Conclusions.

References

R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 407-408; H. L. Hollingworth: *Vocational Psychology*, 1920, pp. 1-56; H. L. Hollingworth: *Judging Human*

Character, 1922, pp. 33-44, 161-200; J. Jastrow: *Fact and Fable in Psychology,* 1900, pp. 21-26, 236-274; J. Jastrow: *Psychology of Conviction,* 1918, pp. 1-74, 128-172; R. Pintner: *Psychological Review,* Vol. 25, 1918, pp. 286-296; P. C. Gaskill, N. Fenton, and J. P. Porter: "Judging the Intelligence of Boys from Their Photographs," *Jour. Appl. Psychol.,* 1927, Vol. 11, pp. 394-403.

CHAPTER XXXVII

Experiment No. 31

JUDGMENT OF EMOTION FROM PHOTOGRAPHS

Even though it appears doubtful that relatively permanent mental characteristics can ever be judged with perfect accuracy from facial contours and expressions, it seems more probable that temporary mental states which involve rather definite movements of the facial muscles may be so judged. That the difficulties of even the latter sort of judgment are not small should be clear from the following experiment.

The purpose of this experiment is to determine how accurately one is able to judge emotional attitudes from photographs.

Apparatus.—Twenty-four photographs, representing emotional attitudes.

Method.—The photographs were posed to represent the following mental states:

1. Interest in a child
2. Religious feeling
3. Horror
4. Agreeable surprise
5. Despair
6. Sneering
7. Breathless interest
8. Strong suspicion
9. Fear
10. Pity
11. Vanity
12. Modesty
13. Laughter
14. Determination
15. Hate
16. Mental multiplication
17. Faint suspicion
18. Righteous anger

19. Sympathy
20. Contempt
21. Attention to an object
22. Physical pain
23. Rage
24. Disgust

Choose a photograph to match each emotion listed. Choose each photograph once only, and when best suited with your matches, report results to the instructor for the class record.

Copy the class record, and get from the instructor a list of the emotions actually intended by the sitter.

Results and Discussion.

(1) What percentage of your judgments agree with the state intended?

(2) From the class record on the blackboard determine for each photograph what number (emotion) is most frequently assigned. What percentage of the most frequently assigned emotions agree with the intended?

(3) What percentage of your assigned emotions agree with these most frequently assigned by the class as a whole? Do you agree better with the intended emotions or with the judgments of emotion made by the class? Summarize in general terms what the above results tend to show.

(4) Describe your difficulties in judging. Offer suggestions for improving the experiment.

(5) What biological and physiological and psychological theories have been given to account for the particular expressions which betray emotion? (See references below.)

(6) Summarize experimental results on the accuracy of judging emotional attitudes from facial expressions

and compare with results obtained in this experiment. (See Dashiell, pp. 427-431.)

Summary and Conclusions.

References

R. S. Woodworth: *Psychology*, 1929 (Rev. Ed.), pp. 285-291; E. B. Titchener: *Text Book of Psychology*, 1911, pp. 484-489; C. Darwin: *The Expression of Emotion in Man and Animals*, 1872; A. M. Feleky: *Psychological Review*, Vol. 21, 1914, pp. 33-41; J. F. Dashiell: *Fundamentals of Objective Psychology*, 1928, pp. 427-431.

CHAPTER XXXVIII

Experiment No. 32

THE MEASUREMENT OF ART TALENT

(An Art Judgment Test) *

Obviously there are several variables involved in artistic talent. After making a study of the methods of the great masters in art, N. C. Meier has concluded that success in graphic art is closely correlated with *æsthetic sensitivity*. He has found that this variable is well suited to laboratory treatment and is of crucial prognostic value.† His studies led to construction of the Meier-Seashore Art Judgment Test as a measure of æsthetic sensitivity. In this test works of recognized artists are utilized as test material. He assumes that "the constant in art" lies in "general principles or qualities of balance, harmony, rhythm and their variations." For each comparison in the test an art work of a master was reproduced twice; once as an exact copy of the original and once with some element (i.e., perspec-

* An experiment on poetry judging may be substituted for the one offered here and the results treated in a manner similar to those described here. For such an experiment use "Exercises in Judging Poetry," Series X and Y (together) by A. Abbott and M. R. Trabue, published by Bureau of Publications, Teachers College, Columbia University, New York City. (Permission for use of test given by Professors M. R. Trabue and A. Abbott.)

In case there is time for both the art judgment and the poetry judgment tests the relation between the two types of artistic judgment can be determined.

† The instructor should expand this introduction. See references.

tive, arrangement of principal object, etc.), altered so as to destroy the balance, rhythm or unity. The subject is to express a preference for the right or left picture after reading a printed statement of what has been altered in a given pair of pictures and after studying them. For a group of 1,081 subjects the scores grouped themselves in approximately a normal distribution. With subgroups the distribution of scores show overlapping for all classes tested. However, the scores made by faculty members in departments of art tend to be grouped in the upper range of scores although some of the other subjects, even though untrained in art work, obtain high scores.*

The purpose of this experiment is to measure the artistic talent of individuals in a college class, to note (1) sex differences in scores, (2) relation of art talent to intelligence, and (3) the relation of art talent to training in that field.

Apparatus.—A copy of the Meier-Seashore Art Judgment Test † and response blank (see *Notes for Instructors*) for each student.

Method.—The instructor acts as experimenter and the class as subjects. Each student computes his total art training score by adding the scores under each of the following five heads:

(1) A score of 1 for total elementary school drawing ——— (= score).

(2) Number of credits (class hours per week for term or quarter) in high school or college art (freehand drawing, mechanical drawing, design, sketching, still life,

* Consult the Meier-Seashore sheet of directions for temporary norms and interpretations of scores.

† Used with the permission of Professor N. C. Meier.

painting, modeling, æsthetics, history of art, etc.) ÷ 3 ——— (= score).

(3) Approximate number of visits to art museum per year for the past four years. (This includes visits to private collections or to galleries in art shops) ——— (= score).

(4) Number of art magazines you read regularly ——— (= score).

Add above units for total art training score. Total score = ———.

After this total score has been determined give the art judgment test to the class according to directions on the record sheet. Have the students exchange blanks and score the test by reference to the correct answers read by the instructor. The number of correct responses is divided by 125 to obtain the per cent right. The art judgment scores (per cent right), the art training scores, and the Alpha intelligence test scores (from Experiment No. 22) are listed on the blackboard in one table for the men and those for the women in another. The class material is then copied for use in computations by individuals.

Results and Discussion.

(1) Compute the sex difference in art Judgment and the reliability of the difference. Interpret your findings.

(2) What is the relation between art training and art talent scores? Explain fully.

(Alternate for question 2: Select for one sub-group those students whose scores make up the highest 25 per cent of the total group; for another sub-group the lowest 25 per cent. Compare these two sub-groups with each other in as many ways as you can and interpret your findings.)

(3) What is the relation between intelligence and art talent scores? Explain fully.

(4) Discuss the possible applications and value of this art judgment test.

(5) How could the reliability of this test be computed?

References

Norman C. Meier: Æsthetic Judgment as a Measure of Art Talent, *University of Iowa Studies,* Series on Aims and Progress of Research, Vol. I, No. 19, 30 pp; Can Art Talent be Discovered by Test Devices? *23rd Ann. Rpt.,* Western Arts Assn., 1927, pp. 74-79; Special Artistic Talent, *Psychol. Bull.,* 1928, Vol. 25, pp. 265-271; A Measure of Art Talent, *Psychol. Monog.,* 1928, Vol. 39, No. 2, pp. 184-199.

APPENDIX I

STATISTICAL TABLES

Numerical Table—Uses

The use of this table will aid the student in any work which involves the squaring of numbers or the extraction of square roots. In this course in experimental psychology the table will greatly facilitate the computation of standard deviations, coefficients of correlation, and standard errors, and probable errors of averages and differences. In the table the column headed n contains the numbers to be squared, column n^2 the squares of the numbers, column $\sqrt{n}$ the square roots of the numbers and column $\sqrt{10n}$ the square roots of 10 times the numbers.

For squares of numbers 11 to 99, look up in n^2 — column the squares of 1.10, 1.20, 1.30, etc., omitting the two zeros and the decimal point from the squares; thus from $1.10^2 = 1.2100$, $11^2 = 121$; $1.20^2 = 1.4400$, $12^2 = 144$; $1.30^2 = 1.6900$, $13^2 = 169$. For squares of numbers 100 to 1000 use the first two columns of the table omitting decimal points; thus $112^2 = 12544$ (from $1.12^2 = 1.2544$). Where the decimal point in the desired number is in a different position than that given in the n — column, shift the point in both the n — column and in the n^2 — column; thus $10.1^2 = 102.01$ (from $1.01^2 = 1.0201$). In employing this table bear in mind that $1.5 = 1.50$, $2.5 = 2.50$, etc. The following examples illustrate how to use the table to obtain square roots:

$\sqrt{3} = \sqrt{3.00} = 1.73.$
$\sqrt{1.06} = 1.03.$
$\sqrt{10.6} = 3.25$ (in column headed $\sqrt{10n}$.).
$\sqrt{11.47} = \sqrt{11.5}$ (approximately) $= 3.39.$
$\sqrt{118.94} = \sqrt{119}$ (approximately) $= 10.9.$

$\sqrt{119}$ is also expressed $\sqrt{1,19.00}$, the $\sqrt{n}$ — column yields 1.09 and the decimal point is moved one place to the right giving 10.9. (Thus moving the decimal point one place to the right in $\sqrt{n}$ — column yields the square roots of 100n. For all work in the elementary psychological laboratory no error of appreciable size is involved by saying 118.94 is equal to 119, etc., when square roots are to be extracted. Thus $\sqrt{118} = 10.9$, $\sqrt{119} = 10.9$, and $\sqrt{120} = 10.9$. There usually occurs no change in the first decimal place of the square roots to accompany small changes in n as illustrated in the above examples.

NUMERICAL TABLE

TABLE 14

SQUARES AND SQUARE ROOTS OF WHOLE NUMBERS AND DECIMALS FOR NUMBERS UP TO AND INCLUDING 1000

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
1.00	1.0000	1.00	3.16	1.50	2.2500	1.22	3.87
1.01	1.0201	1.00	3.18	1.51	2.2801	1.23	3.88
1.02	1.0404	1.01	3.19	1.52	2.3104	1.23	3.90
1.03	1.0609	1.01	3.21	1.53	2.3409	1.24	3.91
1.04	1.0816	1.02	3.22	1.54	2.3716	1.24	3.92
1.05	1.1025	1.02	3.24	1.55	2.4025	1.24	3.94
1.06	1.1236	1.03	3.25	1.56	2.4336	1.25	3.95
1.07	1.1449	1.03	3.27	1.57	2.4649	1.25	3.96
1.08	1.1664	1.04	3.29	1.58	2.4964	1.26	3.97
1.09	1.1881	1.04	3.30	1.59	2.5281	1.26	3.99
1.10	1.2100	1.05	3.32	1.60	2.5600	1.26	4.00
1.11	1.2321	1.05	3.33	1.61	2.5921	1.27	4.01
1.12	1.2544	1.06	3.35	1.62	2.6244	1.27	4.02
1.13	1.2769	1.06	3.36	1.63	2.6569	1.28	4.04
1.14	1.2996	1.07	3.38	1.64	2.6896	1.28	4.05
1.15	1.3225	1.07	3.39	1.65	2.7225	1.28	4.06
1.16	1.3456	1.08	3.40	1.66	2.7556	1.29	4.07
1.17	1.3689	1.08	3.42	1.67	2.7889	1.29	4.09
1.18	1.3924	1.09	3.43	1.68	2.8224	1.30	4.10
1.19	1.4161	1.09	3.45	1.69	2.8561	1.30	4.11
1.20	1.4400	1.09	3.46	1.70	2.8900	1.30	4.12
1.21	1.4641	1.10	3.48	1.71	2.9241	1.31	4.13
1.22	1.4884	1.10	3.49	1.72	2.9584	1.31	4.15
1.23	1.5129	1.11	3.51	1.73	2.9929	1.31	4.16
1.24	1.5376	1.11	3.52	1.74	3.0276	1.32	4.17
1.25	1.5625	1.12	3.53	1.75	3.0625	1.32	4.18
1.26	1.5876	1.12	3.55	1.76	3.0976	1.33	4.19
1.27	1.6129	1.13	3.56	1.77	3.1329	1.33	4.21
1.28	1.6384	1.13	3.58	1.78	3.1684	1.33	4.22
1.29	1.6641	1.13	3.59	1.79	3.2041	1.34	4.23
1.30	1.6900	1.14	3.60	1.80	3.2400	1.34	4.24
1.31	1.7161	1.14	3.62	1.81	3.2761	1.34	4.25
1.32	1.7424	1.15	3.63	1.82	3.3124	1.35	4.27
1.33	1.7689	1.15	3.65	1.83	3.3489	1.35	4.28
1.34	1.7956	1.16	3.66	1.84	3.3856	1.36	4.29
1.35	1.8225	1.16	3.67	1.85	3.4225	1.36	4.30
1.36	1.8496	1.17	3.69	1.86	3.4596	1.36	4.31
1.37	1.8769	1.17	3.70	1.87	3.4969	1.37	4.32
1.38	1.9044	1.17	3.71	1.88	3.5344	1.37	4.33
1.39	1.9321	1.18	3.73	1.89	3.5721	1.37	4.35
1.40	1.9600	1.18	3.74	1.90	3.6100	1.38	4.36
1.41	1.9881	1.19	3.75	1.91	3.6481	1.38	4.37
1.42	2.0164	1.19	3.77	1.92	3.6864	1.38	4.38
1.43	2.0449	1.19	3.78	1.93	3.7249	1.39	4.39
1.44	2.0736	1.20	3.79	1.94	3.7636	1.39	4.40
1.45	2.1025	1.20	3.81	1.95	3.8025	1.40	4.41
1.46	2.1316	1.21	3.82	1.96	3.8416	1.40	4.43
1.47	2.1609	1.21	3.83	1.97	3.8809	1.40	4.44
1.48	2.1904	1.22	3.85	1.98	3.9204	1.41	4.45
1.49	2.2201	1.22	3.86	1.99	3.9601	1.41	4.46
1.50	2.2500	1.22	3.87	2.00	4.0000	1.41	4.47

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
2.00	4.0000	1.41	4.47	2.50	6.2500	1.58	5.00
2.01	4.0401	1.42	4.48	2.51	6.3001	1.58	5.01
2.02	4.0804	1.42	4.49	2.52	6.3504	1.59	5.02
2.03	4.1209	1.42	4.50	2.53	6.4009	1.59	5.03
2.04	4.1616	1.43	4.52	2.54	6.4516	1.59	5.04
2.05	4.2025	1.43	4.53	2.55	6.5025	1.60	5.05
2.06	4.2436	1.43	4.54	2.56	6.5536	1.60	5.06
2.07	4.2849	1.44	4.55	2.57	6.6049	1.60	5.07
2.08	4.3264	1.44	4.56	2.58	6.6564	1.61	5.08
2.09	4.3681	1.44	4.57	2.59	6.7081	1.61	5.09
2.10	4.4100	1.45	4.58	2.60	6.7600	1.61	5.10
2.11	4.4521	1.45	4.59	2.61	6.8121	1.61	5.11
2.12	4.4944	1.46	4.60	2.62	6.8644	1.62	5.12
2.13	4.5369	1.46	4.61	2.63	6.9169	1.62	5.13
2.14	4.5796	1.46	4.63	2.64	6.9696	1.62	5.14
2.15	4.6225	1.47	4.64	2.65	7.0225	1.63	5.15
2.16	4.6656	1.47	4.65	2.66	7.0756	1.63	5.16
2.17	4.7089	1.47	4.66	2.67	7.1289	1.63	5.17
2.18	4.7524	1.48	4.67	2.68	7.1824	1.64	5.18
2.19	4.7961	1.48	4.68	2.69	7.2361	1.64	5.19
2.20	4.8400	1.48	4.69	2.70	7.2900	1.64	5.20
2.21	4.8841	1.49	4.70	2.71	7.3441	1.65	5.20
2.22	4.9284	1.49	4.71	2.72	7.3984	1.65	5.21
2.23	4.9729	1.49	4.72	2.73	7.4529	1.65	5.22
2.24	5.0176	1.50	4.73	2.74	7.5076	1.65	5.23
2.25	5.0625	1.50	4.74	2.75	7.5625	1.66	5.24
2.26	5.1076	1.50	4.75	2.76	7.6176	1.66	5.25
2.27	5.1529	1.51	4.76	2.77	7.6729	1.66	5.26
2.28	5.1984	1.51	4.77	2.78	7.7284	1.67	5.27
2.29	5.2441	1.51	4.78	2.79	7.7841	1.67	5.28
2.30	5.2900	1.52	4.79	2.80	7.8400	1.67	5.29
2.31	5.3361	1.52	4.81	2.81	7.8961	1.68	5.30
2.32	5.3824	1.52	4.82	2.82	7.9524	1.68	5.31
2.33	5.4289	1.53	4.83	2.83	8.0089	1.68	5.32
2.34	5.4756	1.53	4.84	2.84	8.0656	1.68	5.33
2.35	5.5225	1.53	4.85	2.85	8.1225	1.69	5.34
2.36	5.5696	1.54	4.86	2.86	8.1796	1.69	5.35
2.37	5.6169	1.54	4.87	2.87	8.2369	1.69	5.36
2.38	5.6644	1.54	4.88	2.88	8.2944	1.70	5.37
2.39	5.7121	1.54	4.89	2.89	8.3521	1.70	5.37
2.40	5.7600	1.55	4.90	2.90	8.4100	1.70	5.38
2.41	5.8081	1.55	4.91	2.91	8.4681	1.70	5.39
2.42	5.8564	1.55	4.92	2.92	8.5264	1.71	5.40
2.43	5.9049	1.56	4.93	2.93	8.5849	1.71	5.41
2.44	5.9536	1.56	4.94	2.94	8.6436	1.71	5.42
2.45	6.0025	1.56	4.95	2.95	8.7025	1.72	5.43
2.46	6.0516	1.57	4.96	2.96	8.7616	1.72	5.44
2.47	6.1009	1.57	4.97	2.97	8.8209	1.72	5.45
2.48	6.1504	1.57	4.98	2.98	8.8804	1.73	5.46
2.49	6.2001	1.58	4.99	2.99	8.9401	1.73	5.47
2.50	6.2500	1.58	5.00	3.00	9.0000	1.73	5.48

n	n²	√n	√10n	n	n²	√n	√10n
3.00	**9.0000**	**1.73**	**5.48**	**3.50**	**12.2500**	**1.87**	**5.92**
3.01	9.0601	1.73	5.49	3.51	12.3201	1.87	5.92
3.02	9.1204	1.74	5.49	3.52	12.3904	1.88	5.93
3.03	9.1809	1.74	5.50	3.53	12.4609	1.88	5.94
3.04	9.2416	1.74	5.51	3.54	12.5316	1.88	5.95
3.05	9.3025	1.75	5.52	3.55	12.6025	1.88	5.96
3.06	9.3636	1.75	5.53	3.56	12.6736	1.89	5.97
3.07	9.4249	1.75	5.54	3.57	12.7449	1.89	5.97
3.08	9.4864	1.75	5.55	3.58	12.8164	1.89	5.98
3.09	9.5481	1.76	5.56	3.59	12.8881	1.89	5.99
3.10	9.6100	1.76	5.57	3.60	12.9600	1.90	6.00
3.11	9.6721	1.76	5.58	3.61	13.0321	1.90	6.01
3.12	9.7344	1.77	5.58	3.62	13.1044	1.90	6.02
3.13	9.7969	1.77	5.59	3.63	13.1769	1.90	6.02
3.14	9.8596	1.77	5.60	3.64	13.2496	1.91	6.03
3.15	9.9225	1.77	5.61	3.65	13.3225	1.91	6.04
3.16	9.9856	1.78	5.62	3.66	13.3956	1.91	6.05
3.17	10.0489	1.78	5.63	3.67	13.4689	1.91	6.06
3.18	10.1124	1.78	5.64	3.68	13.5424	1.92	6.07
3.19	10.1761	1.79	5.65	3.69	13.6161	1.92	6.07
3.20	10.2400	1.79	5.66	3.70	13.6900	1.92	6.08
3.21	10.3041	1.79	5.66	3.71	13.7641	1.93	6.09
3.22	10.3684	1.79	5.67	3.72	13.8384	1.93	6.10
3.23	10.4329	1.80	5.68	3.73	13.9129	1.93	6.11
3.24	10.4976	1.80	5.69	3.74	13.9876	1.93	6.11
3.25	10.5625	1.80	5.70	3.75	14.0625	1.94	6.12
3.26	10.6276	1.80	5.71	3.76	14.1376	1.94	6.13
3.27	10.6929	1.81	5.72	3.77	14.2129	1.94	6.14
3.28	10.7584	1.81	5.73	3.78	14.2884	1.94	6.15
3.29	10.8241	1.81	5.73	3.79	14.3641	1.95	6.16
3.30	10.8900	1.82	5.74	3.80	14.4400	1.95	6.16
3.31	10.9561	1.82	5.75	3.81	14.5161	1.95	6.17
3.32	11.0224	1.82	5.76	3.82	14.5924	1.95	6.18
3.33	11.0889	1.82	5.77	3.83	14.6689	1.96	6.19
3.34	11.1556	1.83	5.78	3.84	14.7456	1.96	6.20
3.35	11.2225	1.83	5.79	3.85	14.8225	1.96	6.20
3.36	11.2896	1.83	5.80	3.86	14.8996	1.96	6.21
3.37	11.3569	1.83	5.80	3.87	14.9769	1.97	6.22
3.38	11.4244	1.84	5.81	3.88	15.0544	1.97	6.23
3.39	11.4921	1.84	5.82	3.89	15.1321	1.97	6.24
3.40	11.5600	1.84	5.83	3.90	15.2100	1.97	6.24
3.41	11.6281	1.85	5.84	3.91	15.2881	1.98	6.25
3.42	11.6964	1.85	5.85	3.92	15.3664	1.98	6.26
3.43	11.7649	1.85	5.86	3.93	15.4449	1.98	6.27
3.44	11.8336	1.85	5.86	3.94	15.5236	1.98	6.28
3.45	11.9025	1.86	5.87	3.95	15.6025	1.99	6.28
3.46	11.9716	1.86	5.88	3.96	15.6816	1.99	6.29
3.47	12.0409	1.86	5.89	3.97	15.7609	1.99	6.30
3.48	12.1104	1.86	5.90	3.98	15.8404	1.99	6.31
3.49	12.1801	1.87	5.91	3.99	15.9201	2.00	6.32
3.50	12.2500	1.87	5.92	4.00	16.0000	2.00	6.32

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
4.00	16.0000	2.00	6.32	4.50	20.2500	2.12	6.71
4.01	16.0801	2.00	6.33	4.51	20.3401	2.12	6.71
4.02	16.1604	2.00	6.34	4.52	20.4304	2.13	6.72
4.03	16.2409	2.01	6.35	4.53	20.5209	2.13	6.73
4.04	16.3216	2.01	6.36	4.54	20.6116	2.13	6.74
4.05	16.4025	2.01	6.36	4.55	20.7025	2.13	6.74
4.06	16.4836	2.01	6.37	4.56	20.7936	2.13	6.75
4.07	16.5649	2.02	6.38	4.57	20.8849	2.14	6.76
4.08	16.6464	2.02	6.39	4.58	20.9764	2.14	6.77
4.09	16.7281	2.02	6.39	4.59	21.0681	2.14	6.77
4.10	16.8100	2.02	6.40	4.60	21.1600	2.14	6.78
4.11	16.8921	2.03	6.41	4.61	21.2521	2.15	6.79
4.12	16.9744	2.03	6.42	4.62	21.3444	2.15	6.80
4.13	17.0569	2.03	6.43	4.63	21.4369	2.15	6.80
4.14	17.1396	2.03	6.43	4.64	21.5296	2.15	6.81
4.15	17.2225	2.04	6.44	4.65	21.6225	2.16	6.82
4.16	17.3056	2.04	6.45	4.66	21.7156	2.16	6.83
4.17	17.3889	2.04	6.46	4.67	21.8089	2.16	6.83
4.18	17.4724	2.04	6.46	4.68	21.9024	2.16	6.84
4.19	17.5561	2.05	6.47	4.69	21.9961	2.16	6.85
4.20	17.6400	2.05	6.48	4.70	22.0900	2.17	6.85
4.21	17.7241	2.05	6.49	4.71	22.1841	2.17	6.86
4.22	17.8084	2.05	6.50	4.72	22.2784	2.17	6.87
4.23	17.8929	2.06	6.50	4.73	22.3729	2.17	6.88
4.24	17.9776	2.06	6.51	4.74	22.4676	2.18	6.88
4.25	18.0625	2.06	6.52	4.75	22.5625	2.18	6.89
4.26	18.1476	2.06	6.53	4.76	22.6576	2.18	6.90
4.27	18.2329	2.07	6.53	4.77	22.7529	2.18	6.91
4.28	18.3184	2.07	6.54	4.78	22.8484	2.19	6.91
4.29	18.4041	2.07	6.55	4.79	22.9441	2.19	6.92
4.30	18.4900	2.07	6.56	4.80	23.0400	2.19	6.93
4.31	18.5761	2.08	6.56	4.81	23.1361	2.19	6.93
4.32	18.6624	2.08	6.57	4.82	23.2324	2.19	6.94
4.33	18.7489	2.08	6.58	4.83	23.3289	2.20	6.95
4.34	18.8356	2.08	6.59	4.84	23.4256	2.20	6.96
4.35	18.9225	2.08	6.59	4.85	23.5225	2.20	6.96
4.36	19.0096	2.09	6.60	4.86	23.6196	2.20	6.97
4.37	19.0969	2.09	6.61	4.87	23.7169	2.21	6.98
4.38	19.1844	2.09	6.62	4.88	23.8144	2.21	6.98
4.39	19.2721	2.09	6.62	4.89	23.9121	2.21	6.99
4.40	19.3600	2.10	6.63	4.90	24.0100	2.21	7.00
4.41	19.4481	2.10	6.64	4.91	24.1081	2.21	7.01
4.42	19.5364	2.10	6.65	4.92	24.2064	2.22	7.01
4.43	19.6249	2.10	6.65	4.93	24.3049	2.22	7.02
4.44	19.7136	2.11	6.66	4.94	24.4036	2.22	7.03
4.45	19.8025	2.11	6.67	4.95	24.5025	2.22	7.03
4.46	19.8916	2.11	6.68	4.96	24.6016	2.23	7.04
4.47	19.9809	2.11	6.68	4.97	24.7009	2.23	7.05
4.48	20.0704	2.12	6.69	4.98	24.8004	2.23	7.06
4.49	20.1601	2.12	6.70	4.99	24.9001	2.23	7.06
4.50	20.2500	2.12	6.71	5.00	25.0000	2.24	7.07

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
5.00	25.0000	2.24	7.07	5.50	30.2500	2.34	7.42
5.01	25.1001	2.24	7.08	5.51	30.3601	2.35	7.42
5.02	25.2004	2.24	7.08	5.52	30.4704	2.35	7.43
5.03	25.3009	2.24	7.09	5.53	30.5809	2.35	7.44
5.04	25.4016	2.24	7.10	5.54	30.6916	2.35	7.44
5.05	25.5025	2.25	7.11	5.55	30.8025	2.35	7.45
5.06	25.6036	2.25	7.11	5.56	30.9136	2.36	7.46
5.07	25.7049	2.25	7.12	5.57	31.0249	2.36	7.46
5.08	25.8064	2.25	7.13	5.58	31.1364	2.36	7.47
5.09	25.9081	2.26	7.13	5.59	31.2481	2.36	7.48
5.10	26.0100	2.26	7.14	5.60	31.3600	2.37	7.48
5.11	26.1121	2.26	7.15	5.61	31.4721	2.37	7.49
5.12	26.2144	2.26	7.15	5.62	31.5844	2.37	7.50
5.13	26.3169	2.26	7.16	5.63	31.6969	2.37	7.50
5.14	26.4196	2.27	7.17	5.64	31.8096	2.37	7.51
5.15	26.5225	2.27	7.18	5.65	31.9225	2.38	7.52
5.16	26.6256	2.27	7.18	5.66	32.0356	2.38	7.52
5.17	26.7289	2.27	7.19	5.67	32.1489	2.38	7.53
5.18	26.8324	2.27	7.20	5.68	32.2624	2.38	7.54
5.19	26.9361	2.28	7.20	5.69	32.3761	2.38	7.54
5.20	27.0400	2.28	7.21	5.70	32.4900	2.39	7.55
5.21	27.1441	2.28	7.22	5.71	32.6041	2.39	7.56
5.22	27.2484	2.28	7.22	5.72	32.7184	2.39	7.56
5.23	27.3529	2.29	7.23	5.73	32.8329	2.39	7.57
5.24	27.4576	2.29	7.24	5.74	32.9476	2.39	7.58
5.25	27.5625	2.29	7.24	5.75	33.0625	2.40	7.58
5.26	27.6676	2.29	7.25	5.76	33.1776	2.40	7.59
5.27	27.7729	2.29	7.26	5.77	33.2929	2.40	7.60
5.28	27.8784	2.30	7.27	5.78	33.4084	2.40	7.60
5.29	27.9841	2.30	7.27	5.79	33.5241	2.41	7.61
5.30	28.0900	2.30	7.28	5.80	33.6400	2.41	7.61
5.31	28.1961	2.30	7.29	5.81	33.7561	2.41	7.62
5.32	28.3024	2.31	7.29	5.82	33.8724	2.41	7.63
5.33	28.4089	2.31	7.30	5.83	33.9889	2.41	7.63
5.34	28.5156	2.31	7.31	5.84	34.1056	2.42	7.64
5.35	28.6225	2.31	7.31	5.85	34.2225	2.42	7.65
5.36	28.7296	2.31	7.32	5.86	34.3396	2.42	7.65
5.37	28.8369	2.32	7.33	5.87	34.4569	2.42	7.66
5.38	28.9444	2.32	7.33	5.88	34.5744	2.42	7.67
5.39	29.0521	2.32	7.34	5.89	34.6921	2.43	7.67
5.40	29.1600	2.32	7.35	5.90	34.8100	2.43	7.68
5.41	29.2681	2.32	7.35	5.91	34.9281	2.43	7.69
5.42	29.3764	2.33	7.36	5.92	35.0464	2.43	7.69
5.43	29.4849	2.33	7.37	5.93	35.1649	2.43	7.70
5.44	29.5936	2.33	7.37	5.94	35.2836	2.44	7.71
5.45	29.7025	2.33	7.38	5.95	35.4025	2.44	7.71
5.46	29.8116	2.34	7.39	5.96	35.5216	2.44	7.72
5.47	29.9209	2.34	7.39	5.97	35.6409	2.44	7.73
5.48	30.0304	2.34	7.40	5.98	35.7604	2.44	7.73
5.49	30.1401	2.34	7.41	5.99	35.8801	2.45	7.74
5.50	30.2500	2.34	7.42	6.00	36.0000	2.45	7.74

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
6.00	36.0000	2.45	7.74	6.50	42.2500	2.55	8.06
6.01	36.1201	2.45	7.75	6.51	42.3801	2.55	8.07
6.02	36.2404	2.45	7.76	6.52	42.5104	2.55	8.07
6.03	36.3609	2.45	7.76	6.53	42.6409	2.55	8.08
6.04	36.4816	2.46	7.77	6.54	42.7716	2.56	8.09
6.05	36.6025	2.46	7.78	6.55	42.9025	2.56	8.09
6.06	36.7236	2.46	7.78	6.56	43.0336	2.56	8.10
6.07	36.8449	2.46	7.79	6.57	43.1649	2.56	8.10
6.08	36.9664	2.46	7.80	6.58	43.2964	2.56	8.11
6.09	37.0881	2.47	7.80	6.59	43.4281	2.57	8.12
6.10	37.2100	2.47	7.81	6.60	43.5600	2.57	8.12
6.11	37.3321	2.47	7.82	6.61	43.6921	2.57	8.13
6.12	37.4544	2.47	7.82	6.62	43.8244	2.57	8.14
6.13	37.5769	2.47	7.83	6.63	43.9569	2.57	8.14
6.14	37.6996	2.48	7.83	6.64	44.0896	2.58	8.15
6.15	37.8225	2.48	7.84	6.65	44.2225	2.58	8.15
6.16	37.9456	2.48	7.85	6.66	44.3556	2.58	8.16
6.17	38.0689	2.48	7.85	6.67	44.4889	2.58	8.17
6.18	38.1924	2.48	7.86	6.68	44.6224	2.58	8.17
6.19	38.3161	2.49	7.87	6.69	44.7561	2.59	8.18
6.20	38.4400	2.49	7.87	6.70	44.8900	2.59	8.18
6.21	38.5641	2.49	7.88	6.71	45.0241	2.59	8.19
6.22	38.6884	2.49	7.89	6.72	45.1584	2.59	8.20
6.23	38.8129	2.50	7.89	6.73	45.2929	2.59	8.20
6.24	38.9376	2.50	7.90	6.74	45.4276	2.60	8.21
6.25	39.0625	2.50	7.90	6.75	45.5625	2.60	8.21
6.26	39.1876	2.50	7.91	6.76	45.6976	2.60	8.22
6.27	39.3129	2.50	7.92	6.77	45.8329	2.60	8.23
6.28	39.4384	2.50	7.92	6.78	45.9684	2.60	8.23
6.29	39.5641	2.51	7.93	6.79	46.1041	2.60	8.24
6.30	39.6900	2.51	7.94	6.80	46.2400	2.61	8.25
6.31	39.8161	2.51	7.94	6.81	46.3761	2.61	8.25
6.32	39.9424	2.51	7.95	6.82	46.5124	2.61	8.26
6.33	40.0689	2.51	7.96	6.83	46.6489	2.61	8.26
6.34	40.1956	2.52	7.96	6.84	46.7856	2.61	8.27
6.35	40.3225	2.52	7.97	6.85	46.9225	2.62	8.28
6.36	40.4496	2.52	7.97	6.86	47.0596	2.62	8.28
6.37	40.5769	2.52	7.98	6.87	47.1969	2.62	8.29
6.38	40.7044	2.52	7.99	6.88	47.3344	2.62	8.29
6.39	40.8321	2.53	7.99	6.89	47.4721	2.62	8.30
6.40	40.9600	2.53	8.00	6.90	47.6100	2.63	8.31
6.41	41.0881	2.53	8.01	6.91	47.7481	2.63	8.31
6.42	41.2164	2.53	8.01	6.92	47.8864	2.63	8.32
6.43	41.3449	2.53	8.02	6.93	48.0249	2.63	8.32
6.44	41.4736	2.54	8.02	6.94	48.1636	2.63	8.33
6.45	41.6025	2.54	8.03	6.95	48.3025	2.64	8.34
6.46	41.7316	2.54	8.04	6.96	48.4416	2.64	8.34
6.47	41.8609	2.54	8.04	6.97	48.5809	2.64	8.35
6.48	41.9904	2.54	8.05	6.98	48.7204	2.64	8.35
6.49	42.1201	2.55	8.06	6.99	48.8601	2.64	8.36
6.50	42.2500	2.55	8.06	7.00	49.0000	2.64	8.37

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
7.00	49.0000	2.64	8.37	7.50	56.2500	2.74	8.66
7.01	49.1401	2.65	8.37	7.51	56.4001	2.74	8.67
7.02	49.2804	2.65	8.38	7.52	56.5504	2.74	8.67
7.03	49.4209	2.65	8.38	7.53	56.7009	2.74	8.68
7.04	49.5616	2.65	8.39	7.54	56.8516	2.74	8.68
7.05	49.7025	2.65	8.40	7.55	57.0025	2.75	8.69
7.06	49.8436	2.66	8.40	7.56	57.1536	2.75	8.69
7.07	49.9849	2.66	8.41	7.57	57.3049	2.75	8.70
7.08	50.1264	2.66	8.41	7.58	57.4564	2.75	8.71
7.09	50.2681	2.66	8.42	7.59	57.6081	2.75	8.71
7.10	50.4100	2.66	8.43	7.60	57.7600	2.76	8.72
7.11	50.5521	2.67	8.43	7.61	57.9121	2.76	8.72
7.12	50.6944	2.67	8.44	7.62	58.0644	2.76	8.73
7.13	50.8369	2.67	8.44	7.63	58.2169	2.76	8.73
7.14	50.9796	2.67	8.45	7.64	58.3696	2.76	8.74
7.15	51.1225	2.67	8.45	7.65	58.5225	2.76	8.75
7.16	51.2656	2.67	8.46	7.66	58.6756	2.77	8.75
7.17	51.4089	2.68	8.47	7.67	58.8289	2.77	8.76
7.18	51.5524	2.68	8.47	7.68	58.9824	2.77	8.76
7.19	51.6961	2.68	8.48	7.69	59.1361	2.77	8.77
7.20	51.8400	2.68	8.48	7.70	59.2900	2.77	8.77
7.21	51.9841	2.68	8.49	7.71	59.4441	2.78	8.78
7.22	52.1284	2.69	8.50	7.72	59.5984	2.78	8.79
7.23	52.2729	2.69	8.50	7.73	59.7529	2.78	8.79
7.24	52.4176	2.69	8.51	7.74	59.9076	2.78	8.80
7.25	52.5625	2.69	8.51	7.75	60.0625	2.78	8.80
7.26	52.7076	2.69	8.52	7.76	60.2176	2.78	8.81
7.27	52.8529	2.70	8.53	7.77	60.3729	2.79	8.81
7.28	52.9984	2.70	8.53	7.78	60.5284	2.79	8.82
7.29	53.1441	2.70	8.54	7.79	60.6841	2.79	8.83
7.30	53.2900	2.70	8.54	7.80	60.8400	2.79	8.83
7.31	53.4361	2.70	8.55	7.81	60.9961	2.79	8.84
7.32	53.5824	2.70	8.55	7.82	61.1524	2.80	8.84
7.33	53.7289	2.71	8.56	7.83	61.3089	2.80	8.85
7.34	53.8756	2.71	8.57	7.84	61.4656	2.80	8.85
7.35	54.0225	2.71	8.57	7.85	61.6225	2.80	8.86
7.36	54.1696	2.71	8.58	7.86	61.7796	2.80	8.86
7.37	54.3169	2.71	8.58	7.87	61.9369	2.80	8.87
7.38	54.4644	2.72	8.59	7.88	62.0944	2.81	8.88
7.39	54.6121	2.72	8.60	7.89	62.2521	2.81	8.88
7.40	54.7600	2.72	8.60	7.90	62.4100	2.81	8.89
7.41	54.9081	2.72	8.61	7.91	62.5681	2.81	8.89
7.42	55.0564	2.72	8.61	7.92	62.7264	2.81	8.90
7.43	55.2049	2.72	8.62	7.93	62.8849	2.82	8.90
7.44	55.3536	2.73	8.62	7.94	63.0436	2.82	8.91
7.45	55.5025	2.73	8.63	7.95	63.2025	2.82	8.92
7.46	55.6516	2.73	8.64	7.96	63.3616	2.82	8.92
7.47	55.8009	2.73	8.64	7.97	63.5209	2.82	8.93
7.48	55.9504	2.73	8.65	7.98	63.6804	2.82	8.93
7.49	56.1001	2.74	8.65	7.99	63.8401	2.83	8.94
7.50	56.2500	2.74	8.66	8.00	64.0000	2.83	8.94

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$
8.00	64.0000	2.83	8.94	8.50	72.2500	2.91	9.22
8.01	64.1601	2.83	8.95	8.51	72.4201	2.92	9.22
8.02	64.3204	2.83	8.95	8.52	72.5904	2.92	9.23
8.03	64.4809	2.83	8.96	8.53	72.7609	2.92	9.23
8.04	64.6416	2.83	8.97	8.54	72.9316	2.92	9.24
8.05	64.8025	2.84	8.97	8.55	73.1025	2.92	9.25
8.06	64.9636	2.84	8.98	8.56	73.2736	2.92	9.25
8.07	65.1249	2.84	8.98	8.57	73.4449	2.93	9.26
8.08	65.2864	2.84	8.99	8.58	73.6164	2.93	9.26
8.09	65.4481	2.84	8.99	8.59	73.7881	2.93	9.27
8.10	65.6100	2.85	9.00	8.60	73.9600	2.93	9.27
8.11	65.7721	2.85	9.00	8.61	74.1321	2.93	9.28
8.12	65.9344	2.85	9.01	8.62	74.3044	2.93	9.28
8.13	66.0969	2.85	9.02	8.63	74.4769	2.94	9.29
8.14	66.2596	2.85	9.02	8.64	74.6496	2.94	9.29
8.15	66.4225	2.85	9.03	8.65	74.8225	2.94	9.30
8.16	66.5856	2.86	9.03	8.66	74.9956	2.94	9.30
8.17	66.7489	2.86	9.04	8.67	75.1689	2.94	9.31
8.18	66.9124	2.86	9.04	8.68	75.3424	2.95	9.32
8.19	67.0761	2.86	9.05	8.69	75.5161	2.95	9.32
8.20	67.2400	2.86	9.05	8.70	75.6900	2.95	9.33
8.21	67.4041	2.86	9.06	8.71	75.8640	2.95	9.33
8.22	67.5684	2.87	9.07	8.72	76.0385	2.95	9.34
8.23	67.7329	2.87	9.07	8.73	76.2129	2.95	9.34
8.24	67.8976	2.87	9.08	8.74	76.3876	2.96	9.35
8.25	68.0625	2.87	9.08	8.75	76.5625	2.96	9.35
8.26	68.2276	2.87	9.09	8.76	76.7376	2.96	9.36
8.27	68.3929	2.87	9.09	8.77	76.9129	2.96	9.36
8.28	68.5584	2.88	9.10	8.78	77.0884	2.96	9.37
8.29	68.7241	2.88	9.10	8.79	77.2641	2.96	9.37
8.30	68.8900	2.88	9.11	8.80	77.4400	2.97	9.38
8.31	69.0561	2.88	9.11	8.81	77.6161	2.97	9.39
8.32	69.2224	2.88	9.12	8.82	77.7924	2.97	9.39
8.33	69.3889	2.89	9.13	8.83	77.9689	2.97	9.40
8.34	69.5556	2.89	9.13	8.84	78.1456	2.97	9.40
8.35	69.7225	2.89	9.14	8.85	78.3225	2.97	9.41
8.36	69.8896	2.89	9.14	8.86	78.4996	2.98	9.41
8.37	70.0569	2.89	9.15	8.87	78.6769	2.98	9.42
8.38	70.2244	2.89	9.15	8.88	78.8544	2.98	9.42
8.39	70.3921	2.90	9.16	8.89	79.0321	2.98	9.43
8.40	70.5600	2.90	9.16	8.90	79.2100	2.98	9.43
8.41	70.7281	2.90	9.17	8.91	79.3881	2.98	9.44
8.42	70.8964	2.90	9.18	8.92	79.5664	2.99	9.44
8.43	71.0649	2.90	9.18	8.93	79.7449	2.99	9.45
8.44	71.2336	2.90	9.19	8.94	79.9236	2.99	9.45
8.45	71.4025	2.91	9.19	8.95	80.1025	2.99	9.46
8.46	71.5716	2.91	9.20	8.96	80.2816	2.99	9.46
8.47	71.7409	2.91	9.20	8.97	80.4609	2.99	9.47
8.48	71.9104	2.91	9.21	8.98	80.6404	3.00	9.48
8.49	72.0801	2.91	9.21	8.99	80.8201	3.00	9.48
8.50	72.2500	2.91	9.22	9.00	81.0000	3.00	9.49

n	n^2	$\sqrt{n}$	$\sqrt{10n}$
9.00	81.0000	3.00	9.49
9.01	81.1801	3.00	9.49
9.02	81.3604	3.00	9.50
9.03	81.5409	3.00	9.50
9.04	81.7216	3.01	9.51
9.05	81.9025	3.01	9.51
9.06	82.0836	3.01	9.52
9.07	82.2649	3.01	9.52
9.08	82.4464	3.01	9.53
9.09	82.6281	3.01	9.53
9.10	82.8100	3.02	9.54
9.11	82.9921	3.02	9.54
9.12	83.1744	3.02	9.55
9.13	83.3569	3.02	9.55
9.14	83.5396	3.02	9.56
9.15	83.7225	3.02	9.56
9.16	83.9056	3.03	9.57
9.17	84.0889	3.03	9.58
9.18	84.2724	3.03	9.58
9.19	84.4561	3.03	9.59
9.20	84.6400	3.03	9.59
9.21	84.8241	3.03	9.60
9.22	85.0084	3.04	9.60
9.23	85.1929	3.04	9.61
9.24	85.3776	3.04	9.61
9.25	85.5625	3.04	9.62
9.26	85.7476	3.04	9.62
9.27	85.9329	3.04	9.63
9.28	86.1184	3.05	9.63
9.29	86.3041	3.05	9.64
9.30	86.4900	3.05	9.64
9.31	86.6761	3.05	9.65
9.32	86.8624	3.05	9.65
9.33	87.0489	3.05	9.66
9.34	87.2356	3.06	9.66
9.35	87.4225	3.06	9.67
9.36	87.6096	3.06	9.67
9.37	87.7969	3.06	9.68
9.38	87.9844	3.06	9.68
9.39	88.1721	3.06	9.69
9.40	88.3600	3.06	9.69
9.41	88.5481	3.07	9.70
9.42	88.7364	3.07	9.70
9.43	88.9249	3.07	9.71
9.44	89.1136	3.07	9.71
9.45	89.3025	3.07	9.72
9.46	89.4916	3.07	9.73
9.47	89.6809	3.08	9.73
9.48	89.8704	3.08	9.74
9.49	90.0601	3.08	9.74
9.50	90.2500	3.08	9.75

n	n^2	$\sqrt{n}$	$\sqrt{10n}$
9.50	90.2500	3.08	9.75
9.51	90.4401	3.08	9.75
9.52	90.6304	3.08	9.76
9.53	90.8209	3.09	9.76
9.54	91.0116	3.09	9.77
9.55	91.2025	3.09	9.77
9.56	91.3936	3.09	9.78
9.57	91.5849	3.09	9.78
9.58	91.7764	3.09	9.79
9.59	91.9681	3.10	9.79
9.60	92.1600	3.10	9.80
9.61	92.3521	3.10	9.80
9.62	92.5444	3.10	9.81
9.63	92.7369	3.10	9.81
9.64	92.9296	3.10	9.82
9.65	93.1225	3.11	9.82
9.66	93.3156	3.11	9.83
9.67	93.5089	3.11	9.83
9.68	93.7024	3.11	9.84
9.69	93.8961	3.11	9.84
9.70	94.0900	3.11	9.85
9.71	94.2841	3.12	9.85
9.72	94.4784	3.12	9.86
9.73	94.6729	3.12	9.86
9.74	94.8676	3.12	9.87
9.75	95.0625	3.12	9.87
9.76	95.2576	3.12	9.88
9.77	95.4529	3.12	9.88
9.78	95.6484	3.13	9.89
9.79	95.8441	3.13	9.89
9.80	96.0400	3.13	9.90
9.81	96.2361	3.13	9.90
9.82	96.4324	3.13	9.91
9.83	96.6289	3.13	9.91
9.84	96.8256	3.14	9.92
9.85	97.0225	3.14	9.92
9.86	97.2196	3.14	9.93
9.87	97.4169	3.14	9.93
9.88	97.6144	3.14	9.94
9.89	97.8121	3.14	9.94
9.90	98.0100	3.15	9.95
9.91	98.2081	3.15	9.95
9.92	98.4064	3.15	9.96
9.93	98.6049	3.15	9.96
9.94	98.8036	3.15	9.97
9.95	99.0025	3.15	9.97
9.96	99.2016	3.15	9.98
9.97	99.4009	3.16	9.98
9.98	99.6004	3.16	9.99
9.99	99.8001	3.16	9.99
10.00	100.0000	3.16	10.00

Instructions for Calculating Rank Difference Coefficients with the Use of Tables *

1. Arrange each of the two series of measurements to be correlated in rank order. Every pair of measurements is thus represented by a pair of ranks—the relative positions assigned those measurements in the two series.

2. Obtain the difference between the two ranks assigned each case.

3. Square each of these differences. The squares are read from the Table 14 of Appendix. Note that the decimals may be omitted from the squares without appreciable error.

4. Add these squares of the differences. This sum (ΣD^2) is to be found in the body of Table 15.

5. Refer in the table to the proper column for the number of cases used (n). Thus, if 21 cases were ranked, use the second column on second page of Table 15. Run down this column until you find the entry nearest the obtained ΣD^2.

6. Read the entry in this same line at either side of the page (first or last column). This entry is ρ the desired coefficient of correlation.

* Table 15 which follows is the first part of "*Tables to Facilitate the Computation of Coefficients of Correlation by the Rank Difference Method*" (J. Appl. Psychol., 1920, Vol. 4, pp. 115-125), and is reprinted with the permission of Professor J. P. Porter, Editor. Copies of the complete table in handy pamphlet form may be purchased from the Journal of Applied Psychology, Ohio University, Athens, Ohio: single copies at 20 cents each, ten or more copies at 15 cents each, and in the case of very large orders a somewhat lower price will be arranged.

Rank in Series I	Rank in Series II	Difference	Difference Squared
1	7	6	36
2	9	7	49
3	2	1	1
4	12	8	64
5	5	0	0
6	3	3	9
7	8	1	1
8	1	7	49
9	15	6	36
10	11	1	1
11	13	2	4
12	4	8	64
13	16	3	9
14	19.5	5.5	30.25
15	6	9	81
16	18	2	4
17	14	3	9
18	10	8	64
19	21	2	4
20	19.5	.5	.25
21	17	4	16
			$\Sigma D^2 = 531.5$

Example: 21 cases are arranged in rank order for two variables; the sum of the squares of the differences in the corresponding ranks is found to be 531 (five-tenths is disregarded).

Refer to the second column on second page of the table, run down to the entry which most nearly equals 531, that is 524, and read beside it the coefficient of correlation, .66.

Note: (a) If the obtained ΣD^2 is nearly midway between two entries, use the ρ falling between the ρ's corresponding to the two entries. (If the ΣD^2 in the example were 540, ρ would be .65.) Correlation coefficients accurate to .01 can be read almost instantly. Closer interpolation is useless.

(b) For ΣD^2 greater than those corresponding to $\rho = 0$ (i.e., in case of negative correlation) subtract the obtained ΣD^2 from the entry in that column corresponding to $\rho = -1.00$. Find the ρ corresponding to this number obtained by subtraction. With a negative sign prefixed, this is the desired ρ. For example, if $\Sigma D^2 = 2430$ and $N = 21$, subtract 2430 from 3080 ($=650$); look up ρ corresponding to 650 in the column headed 21; $\rho = -.58$.

TABLE 15

To Facilitate the Computation of $\dot{\rho}$

Number of Cases Ranked

ρ	11	12	13	14	15	16	17	18	19	20	
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	4	6	7	9	11	14	16	19	23	27	.98
.96	9	11	15	18	22	27	33	39	46	53	.96
.94	13	17	22	27	34	41	49	58	68	80	.94
.92	18	23	29	36	45	54	65	78	91	106	.92
.90	22	29	36	46	56	68	82	97	114	133	.90
.88	26	34	44	55	67	82	98	116	137	160	.88
.86	31	40	51	64	78	95	114	136	160	186	.86
.84	35	46	58	73	90	109	131	155	182	213	.84
.82	40	51	66	82	101	122	147	174	205	239	.82
.80	44	57	73	91	112	136	163	194	228	266	.80
.78	48	63	80	100	123	150	180	213	251	293	.78
.76	53	69	87	109	134	163	196	233	274	319	.76
.74	57	74	95	118	146	177	212	252	296	346	.74
.72	62	80	102	127	157	190	228	271	319	372	.72
.70	66	86	109	137	168	204	245	291	342	399	.70
.68	70	92	116	146	179	218	261	310	365	426	.68
.66	75	97	124	155	190	231	277	329	388	452	.66
.64	79	103	131	164	202	245	294	349	410	479	.64
.62	84	109	138	173	213	258	310	368	433	505	.62
.60	88	114	146	182	224	272	326	388	456	532	.60
.58	92	120	153	191	235	286	343	407	479	559	.58
.56	97	126	160	200	246	299	359	426	502	585	.56
54	101	132	167	209	258	313	375	446	524	612	.54
.52	106	137	175	218	269	326	392	465	547	638	.52

ρ	11	12	13	14	15	16	17	18	19	20	
.50	110	143	182	228	280	340	408	485	570	665	.50
.48	114	149	189	237	291	354	424	504	593	692	.48
.46	119	154	197	246	302	367	441	523	616	718	.46
.44	123	160	204	255	314	381	457	543	638	745	.44
.42	128	166	211	264	325	394	473	562	661	771	.42
.40	132	172	218	273	336	408	490	581	684	798	.40
.38	136	177	226	282	347	422	506	601	707	825	.38
.36	141	183	233	291	358	435	522	620	730	851	.36
.34	145	189	240	300	370	449	539	640	752	878	.34
.32	150	194	248	309	381	462	555	659	775	904	.32
.30	154	200	255	319	392	476	571	678	798	931	.30
.28	158	206	262	328	403	490	588	698	821	958	.28
.26	163	212	269	337	414	503	604	717	844	984	.26
.24	167	217	277	346	426	517	620	736	866	1011	.24
.22	172	223	284	355	437	530	636	756	889	1037	.22
.20	176	229	291	364	448	544	653	775	912	1064	.20
.18	180	235	298	373	459	558	669	795	935	1091	.18
.16	185	240	306	382	470	571	685	814	958	1117	.16
.14	189	246	313	391	482	585	702	833	980	1144	.14
.12	194	252	320	400	493	598	718	853	1003	1170	.12
.10	198	257	328	410	504	612	734	872	1026	1197	.10
.08	202	263	335	419	515	626	751	891	1049	1224	.08
.06	207	269	342	428	526	639	767	911	1072	1250	.06
.04	211	275	349	437	538	653	783	930	1094	1277	.04
.02	216	280	357	446	549	666	800	950	1117	1303	.02
.00	220	286	364	455	560	680	816	969	1140	1330	.00
— 1.00	440	572	728	910	1120	1360	1632	1938	2280	2660	— 1.00

Number of Cases Ranked

ρ	21	22	23	24	25	26	27	28	29	30	
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	31	35	40	46	52	59	66	73	81	90	.98
.96	62	71	81	92	104	117	131	146	162	180	.96
.94	92	106	121	138	156	176	197	219	244	270	.94
.92	123	142	162	184	208	234	262	292	325	360	.92
.90	154	177	202	230	260	293	328	365	406	450	.90
.88	185	213	243	276	312	351	393	438	487	539	.88
.86	216	248	283	322	364	410	459	512	568	629	.86
.84	246	283	324	368	416	468	524	585	650	719	.84
.82	277	319	364	414	468	527	590	658	731	809	.82
.80	308	354	405	460	520	585	655	731	812	899	.80
.78	339	390	445	506	572	644	721	804	893	989	.78
.76	370	425	486	552	624	702	786	877	974	1079	.76
.74	400	460	526	598	676	761	852	950	1056	1169	.74
.72	431	496	567	644	728	819	917	1023	1137	1259	.72
.70	462	531	607	690	780	878	983	1096	1218	1349	.70
.68	493	567	648	736	832	936	1048	1169	1299	1438	.68
.66	524	602	688	782	884	995	1114	1242	1380	1528	.66
.64	554	638	729	828	936	1053	1179	1315	1462	1618	.64
.62	585	673	769	874	988	1112	1245	1389	1543	1708	.62
.60	616	708	810	920	1040	1170	1310	1462	1624	1798	.60
.58	647	744	850	966	1092	1229	1376	1535	1705	1888	.58
.56	678	779	891	1012	1144	1287	1441	1608	1786	1978	.56
.54	708	815	931	1058	1196	1346	1507	1681	1868	2068	.54
.52	739	850	972	1104	1248	1404	1572	1754	1949	2158	.52

ρ	21	22	23	24	25	26	27	28	29	30	
.50	770	886	1012	1150	1300	1463	1638	1827	2030	2248	.50
.48	801	921	1052	1196	1352	1521	1704	1900	2111	2337	.48
.46	832	956	1093	1242	1404	1580	1769	1973	2192	2427	.46
.44	862	992	1133	1288	1456	1638	1835	2046	2274	2517	.44
.42	893	1027	1174	1334	1508	1697	1900	2119	2355	2607	.42
.40	924	1063	1214	1380	1560	1755	1966	2192	2436	2697	.40
.38	955	1098	1255	1426	1612	1814	2031	2265	2517	2787	.38
.36	986	1133	1295	1472	1664	1872	2097	2338	2598	2877	.36
.34	1016	1169	1336	1518	1716	1931	2162	2412	2680	2967	.34
.32	1047	1204	1376	1564	1768	1989	2228	2485	2761	3057	.32
.30	1078	1240	1417	1610	1820	2048	2293	2558	2842	3147	.30
.28	1109	1275	1457	1656	1872	2106	2359	2631	2923	3236	.28
.26	1140	1311	1498	1702	1924	2165	2424	2704	3004	3326	.26
.24	1170	1346	1538	1748	1976	2223	2490	2777	3086	3416	.24
.22	1201	1381	1579	1794	2028	2282	2555	2850	3167	3506	.22
.20	1232	1417	1619	1840	2080	2340	2621	2923	3248	3596	.20
.18	1263	1452	1660	1886	2132	2399	2686	2996	3329	3686	.18
.16	1294	1488	1700	1932	2184	2457	2752	3069	3410	3776	.16
.14	1324	1523	1741	1978	2236	2516	2817	3142	3492	3866	.14
.12	1355	1558	1781	2024	2288	2574	2883	3216	3573	3956	.12
.10	1386	1594	1822	2070	2340	2633	2948	3289	3654	4046	.10
.08	1417	1629	1862	2116	2392	2691	3014	3362	3735	4135	.08
.06	1448	1665	1903	2162	2444	2750	3079	3435	3816	4225	.06
.04	1478	1700	1943	2208	2496	2808	3145	3508	3898	4315	.04
.02	1509	1736	1984	2254	2548	2867	3210	3581	3979	4405	.02
.00	1540	1771	2024	2300	2600	2925	3276	3654	4060	4495	.00
— 1.00	3080	3542	4048	4600	5200	5850	6552	7308	8120	8990	— 1.00

NUMBER OF CASES RANKED

ρ	31	32	33	34	35	36	37	38	39	40	
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	99	109	120	131	143	155	169	183	198	213	.98
.96	198	218	239	262	286	311	337	366	395	426	.96
.94	298	327	359	393	428	466	506	548	593	640	.94
.92	397	436	479	524	571	622	675	731	790	853	.92
.90	496	546	598	655	714	777	844	914	988	1066	.90
.88	595	655	718	785	857	932	1012	1097	1186	1279	.88
.86	694	764	838	916	1000	1088	1181	1279	1383	1492	.86
.84	794	873	957	1047	1142	1243	1350	1462	1581	1706	.84
.82	893	982	1077	1178	1285	1399	1518	1645	1778	1919	.82
.80	992	1091	1197	1309	1428	1554	1687	1828	1976	2132	.80
.78	1091	1200	1316	1440	1571	1709	1856	2011	2174	2345	.78
.76	1190	1309	1436	1571	1714	1865	2025	2193	2371	2558	.76
.74	1290	1419	1556	1702	1856	2020	2193	2376	2569	2772	.74
.72	1389	1528	1676	1833	1999	2176	2362	2559	2766	2985	.72
.70	1488	1637	1795	1964	2142	2331	2531	2742	2964	3198	.70
.68	1587	1746	1915	2094	2285	2486	2700	2924	3162	3411	.68
.66	1686	1855	2035	2225	2428	2642	2868	3107	3359	3624	.66
.64	1786	1964	2154	2356	2570	2797	3037	3290	3557	3838	.64
.62	1885	2073	2274	2487	2713	2953	3206	3473	3754	4051	.62
.60	1984	2182	2394	2618	2856	3108	3374	3657	3952	4264	.60
.58	2083	2292	2513	2749	2999	3263	3543	3838	4150	4477	.58
.56	2182	2401	2633	2880	3142	3419	3712	4021	4347	4690	.56
.54	2282	2510	2753	3011	3284	3574	3881	4204	4545	4904	.54
.52	2381	2619	2872	3142	3427	3730	4049	4387	4742	5117	.52

ρ	31	32	33	34	35	36	37	38	39	40	
.50	2480	2728	2992	3273	3570	3885	4218	4570	4940	5339	.50
.48	2579	2837	3112	3403	3713	4040	4387	4752	5138	5543	.48
.46	2678	2946	3231	3534	3856	4196	4555	4935	5335	5756	.46
.44	2778	3055	3351	3665	3998	4351	4724	5118	5533	5970	.44
.42	2877	3164	3471	3796	4141	4507	4893	5301	5730	6183	.42
.40	2976	3274	3590	3927	4284	4662	5062	5483	5928	6396	.40
.38	3075	3383	3710	4058	4427	4817	5230	5666	6126	6609	.38
.36	3174	3492	3830	4189	4570	4973	5399	5849	6323	6822	.36
.34	3274	3601	3949	4320	4712	5128	5568	6032	6521	7036	.34
.32	3373	3710	4069	4451	4855	5284	5736	6215	6718	7249	.32
.30	3472	3819	4189	4582	4998	5439	5905	6397	6916	7462	.30
.28	3571	3928	4308	4712	5141	5594	6074	6580	7114	7675	.28
.26	3670	4037	4428	4843	5284	5750	6243	6763	7311	7888	.26
.24	3770	4147	4548	4974	5426	5905	6411	6946	7509	8102	.24
.22	3869	4256	4668	5105	5569	6061	6580	7128	7706	8315	.22
.20	3968	4365	4787	5236	5712	6216	6749	7311	7904	8528	.20
.18	4067	4474	4907	5367	5855	6371	6918	7494	8102	8741	.18
.16	4166	4583	5027	5498	5998	6527	7086	7677	8299	8954	.16
.14	4266	4692	5146	5629	6140	6682	7255	7860	8497	9168	.14
.12	4365	4801	5266	5760	6283	6838	7424	8042	8694	9381	.12
.10	4464	4910	5386	5891	6426	6993	7592	8225	8892	9594	.10
.08	4563	5020	5505	6021	6569	7148	7761	8408	9090	9807	.08
.06	4662	5129	5625	6152	6712	7304	7930	8591	9287	10020	.06
.04	4762	5238	5745	6283	6854	7459	8099	8773	9485	10234	.04
.02	4861	5347	5864	6414	6997	7615	8267	8956	9682	10447	.02
.00	4960	5456	5984	6545	7140	7770	8436	9139	9880	10660	.00
− 1.00	9920	10912	11968	13090	14280	15540	16872	18278	19760	21320	— 1.00

NUMBER OF CASES RANKED

ρ	41	42	43	44	45	46	47	48	49	50	
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	230	247	265	284	304	324	346	368	392	417	.98
.96	459	494	530	568	607	649	692	737	784	833	.96
.94	689	740	795	851	911	973	1038	1105	1176	1250	.94
.92	918	987	1060	1135	1214	1297	1384	1474	1568	1666	.92
.90	1148	1234	1324	1419	1518	1622	1730	1842	1960	2083	.90
.88	1378	1481	1589	1703	1822	1946	2076	2211	2352	2499	.88
.86	1607	1728	1854	1987	2125	2270	2421	2579	2744	2916	.86
.84	1837	1975	2119	2270	2429	2594	2767	2948	3136	3332	.84
.82	2066	2221	2384	2554	2732	2919	3113	3316	3528	3749	.82
.80	2296	2468	2649	2838	3036	3243	3459	3685	3920	4165	.80
.78	2526	2715	2914	3122	3340	3567	3805	4053	4312	4582	.78
.76	2755	2962	3179	3406	3643	3892	4151	4422	4704	4998	.76
.74	2985	3209	3443	3689	3947	4216	4497	4790	5096	5415	.74
.72	3214	3455	3708	3973	4250	4540	4843	5159	5488	5831	.72
.70	3444	3702	3973	4257	4554	4865	5189	5527	5880	6248	.70
.68	3674	3949	4238	4541	4858	5189	5535	5896	6272	6664	.68
.66	3903	4196	4503	4825	5161	5513	5881	6264	6664	7081	.66
.64	4133	4443	4768	5108	5465	5837	6227	6633	7056	7497	.64
.62	4362	4690	5033	5392	5768	6162	6572	7001	7448	7914	.62
.60	4592	4936	5298	5676	6072	6486	6918	7370	7840	8330	.60
.58	4822	5183	5562	5960	6376	6810	7264	7738	8232	8747	.58
.56	5051	5430	5827	6244	6679	7135	7610	8107	8624	9163	.56
.54	5281	5677	6092	6527	6983	7459	7956	8475	9016	9580	.54
.52	5510	5924	6357	6811	7286	7783	8302	8844	9408	9996	.52

ρ	41	42	43	44	45	46	47	48	49	50	ρ
.50	5740	6171	6622	7095	7590	8108	8648	9212	9800	10413	.50
.48	5970	6417	6887	7379	7894	8432	8994	9580	10192	10829	.48
.46	6199	6664	7152	7663	8197	8756	9340	9949	10584	11246	.46
.44	6429	6911	7417	7946	8501	9080	9686	10317	10976	11662	.44
.42	6658	7158	7682	8230	8804	9405	10032	10686	11368	12079	.42
.40	6888	7405	7946	8514	9108	9729	10378	11054	11760	12495	.40
.38	7118	7651	8211	8798	9412	10053	10724	11423	12152	12912	.38
.36	7347	7898	8476	9082	9715	10378	11069	11791	12544	13328	.36
.34	7577	8145	8741	9365	10019	10702	11415	12160	12936	13745	.34
.32	7806	8392	9006	9649	10322	11026	11761	12528	13328	14161	.32
.30	8036	8639	9271	9933	10626	11351	12107	12897	13720	14578	.30
.28	8266	8886	9536	10217	10930	11675	12453	13265	14112	14994	.28
.26	8495	9132	9801	10501	11233	11999	12799	13634	14504	15411	.26
.24	8725	9379	10065	10784	11537	12323	13145	14002	14896	15827	.24
.22	8954	9626	10330	11068	11840	12648	13491	14371	15288	16244	.22
.20	9184	9873	10595	11352	12144	12972	13837	14739	15680	16660	.20
.18	9414	10120	10860	11636	12448	13296	14183	15108	16072	17077	.18
.16	9643	10366	11125	11920	12751	13621	14529	15476	16464	17493	.16
.14	9873	10613	11390	12203	13055	13945	14875	15845	16856	17910	.14
.12	10102	10860	11655	12487	13358	14269	15220	16213	17248	18326	.12
.10	10332	11107	11920	12771	13662	14594	15566	16582	17640	18743	.10
.08	10562	11354	12184	13055	13966	14918	15912	16950	18032	19159	.08
.06	10791	11601	12449	13339	14269	15242	16258	17319	18424	19576	.06
.04	11021	11847	12714	13622	14573	15566	16604	17687	18816	19992	.04
.02	11250	12094	12979	13906	14876	15891	16950	18056	19208	20409	.02
.00	11480	12341	13244	14190	15180	16215	17296	18424	19600	20825	.00
— 1.00	22960	24682	26488	28380	30360	32430	34592	36848	39200	41650	— 1.00

TABLE 16

PROBABLE ERRORS OF THE COEFFICIENTS OF CORRELATION FOR VARIOUS NUMBERS OF MEASURES (N) AND FOR VARIOUS VALUES OF ρ. COMPUTED FROM THE FORMULA:

$$PE_\rho = .7063 \frac{1-\rho^2}{\sqrt{N}}$$

(1) ρ	(2) N=10	(3) N=15	(4) N=20	(5) N=25	(6) N=30	(7) N=35	(8) N=40	(9) N=45	(10) N=50
.00	.22	.18	.16	.14	.13	.12	.11	.10	.10
.05	.22	.18	.16	.14	.13	.12	.11	.10	.10
.10	.22	.18	.16	.14	.13	.12	.11	.10	.10
.15	.22	.18	.15	.14	.13	.12	.11	.10	.10
.20	.21	.17	.15	.13	.12	.11	.11	.10	.09
.25	.21	.17	.15	.13	.12	.11	.10	.10	.09
.30	.20	.16	.14	.13	.12	.11	.10	.09	.09
.35	.20	.16	.14	.12	.11	.10	.10	.09	.09
.40	.19	.15	.13	.12	.11	.10	.09	.09	.08
.45	.18	.14	.12	.11	.10	.09	.09	.08	.08
.50	.17	.14	.12	.10	.10	.09	.08	.08	.07
.55	.15	.13	.11	.10	.09	.08	.08	.07	.07
.60	.14	.12	.10	.09	.08	.08	.07	.07	.06
.65	.13	.10	.09	.08	.07	.07	.06	.06	.06
.70	.11	.09	.08	.07	.06	.06	.06	.05	.05
.75	.10	.08	.07	.06	.06	.05	.05	.05	.04
.80	.08	.06	.06	.05	.05	.04	.04	.04	.04
.85	.06	.05	.04	.04	.03	.03	.03	.03	.03
.90	.04	.03	.03	.03	.02	.02	.02	.02	.02
.95	.02	.02	.01	.01	.01	.01	.01	.01	.01

Use Table 16 as follows. Look up the coefficient in the first column headed ρ and opposite this value, in the column headed with the appropriate N read off the desired PE_ρ. Thus where $\rho = .20$ and $N = 25$, look down the first column until you find .20 which is in the fifth row. Follow out on the fifth row to the fifth column ($N = 25$) and read off the desired $PE_\rho = .13$. Again where $\rho = .45$ and $N = 30$, $PE_\rho = .10$. For a coefficient not given in the table, as .675 with $N = 10$, PE_ρ is obtained in column headed $N = 10$, halfway between the values opposite coefficients .65 and .70. Thus $PE_\rho = \frac{(.13 + .11)}{2} = .12$. For practical purposes, where only two decimal places are used, this interpolated value does not involve appreciable error. Similar interpolation may be employed for computing PE_ρ for coefficients with N's not given in the table (i.e., $N = 13$).

APPENDIX II

APPARATUS AND MATERIALS

Instructions for constructing in the laboratory or workshop the simpler pieces of apparatus or attachments for standard pieces referred to in this manual are given here unless information ought not to reach the student's eye before the experiment is completed. In the latter case further detail is given in the special *Notes for Instructors.* The few standard pieces and supplies used must, of course, be obtained from firms dealing in scientific apparatus. C. H. Stoelting & Co., 424 N. Homan Ave., Chicago, Ill., is prepared to supply standard pieces and to undertake the manufacture of special apparatus.

Experiment No. 1. *Individual Differences.*—The materials are described in *Notes for Instructors.*

Experiment No. 2. *The Knee Jerk.*—The reflex hammer is an ordinary ten-cent store tack hammer, with a rubber cane-tip fitted over the smaller (conical) end. The cardboard sheet bearing the scale is about 50 x 35 cm. with a scale in degrees drawn in India ink approximately 43 cm. from the pivot hole. The latter is a metal eyelet. The stand is of the common, flat base type used in chemistry. The oblong base of the stand is 5 x 8 inches with a 19-inch upright rod. To this upright and near the base an iron cross-rod of about the same length is attached by an ordinary right angle clamp. Near the outer end of the cross-rod is fixed a universal clamp which also

grasps the lower edge of the cardboard. Into the second jaw of this clamp goes the ten-penny nail which acts as the pivot. A clamp on the upper part of the vertical rod holds the upper edge of the cardboard rigid.

An ordinary meter stick will serve for the lever. An ordinary rubber band across the cardboard and over the stick will by friction prevent the stick from falling back from the maximum reading. The whole arrangement can readily be taken down and stored compactly or the parts may be used for other experiments.

The shoe-attachment is cut in one piece from a sheet of tin, the two strips which project at right angles then being bent as indicated in Fig. 8. These strips project approximately 9 cm. from the upright portion. The latter is attached by rubber bands so that one strip crosses just in front of the heel. The subject can walk without removing the attachment.

Experiment No. 3. *The Simple Reaction.*—The base of our inexpensive vernier pendulum is the 5 x 8-inch oblong base standard used previously, and described for the knee-jerk register. Two holes about 6 inches apart are drilled through this base near one long edge. Through these holes go screws to hold the wooden strip which bears the clips which serve as keys. This strip is about 7 inches long, 1 inch thick at its two ends, and of a shape indicated in Fig. 9.

The keys are ordinary "Bull-Dog" clips, which are of the "take down" pattern. The jaws are removed from the cylindrical steel spring, and a second hole is drilled in one jaw, this jaw reinserted in the spring, and then fastened to the wooden base by two small screws. The second jaw is then inserted.

The horizontal arm is a hexagonal rod of brass $5\frac{1}{2}$ inches long, and is held by a universal clamp. The rod is drilled with 4 holes for the threads at distances $1\frac{1}{2}$, $3\frac{1}{4}$, and $4\frac{3}{4}$ inches from the fixed end. About $\frac{1}{2}$ inch beyond the first and third holes, two additional larger holes are drilled to hold the copper rivets and washers which serve as bobbins.

The pendulum bobs are sawed $\frac{1}{2}$ inch thick from $1\frac{1}{8}$-inch round brass rod. Small holes 90° from each other are drilled near the edge to take the thread and the wire which catches in the clip. All thread holes are slightly countersunk to avoid sharp edges, which would cut or wear the threads.

Experiment No. 4. *Set and Complex Behavior.*—The paragraphs containing hidden words have been used in mimeographed form. A copy in which the hidden words are underlined is given in the *Notes for Instructors.*

We have used a separate 3 x 5 card for each of our anagrams and skeleton-words and numbered each card to correspond to the numbers in the lists. A more convenient method may be to print or mimeograph the lists, with wide spacing, and then to present the desired word through a window in a sheet of cardboard which covers the remaining words on the page. The lists used, and the key lists are given in *Notes for Instructors.*

Experiment No. 5. *Cutaneous Sensations.*—The stamp makes a map 20 x 20 mm. and is divided into one hundred 2 x 2 mm. squares. The two central cross-lines are widened to facilitate orientation.

The temperature cylinders are turned from round brass rods, 4 inches in length, and $\frac{3}{8}$ inch in diameter.

The head is turned down to a circle of 5 mm. diameter, the other end to a blunt point. Forty-penny wire nails make fairly good substitutes if the point is filed down and smoothed so that it does not scratch the skin.

The stimulator for pressure and pain is made by cutting a horse hair about 4 inches long squarely across at one end and obliquely across at the other, and inserting the hair in a slit cut into one side of a cork.

Experiment No. 6. *Color Mixture.*—Discs about 8 and about 4 inches in diameter, cut from the Milton-Bradley colored papers, are used except in case of blue and yellow. For the latter the Hering complementary papers were not procurable, and the Department of Psychology at the University of Minnesota has had blue and yellow papers prepared which appear quite complementary under normal working conditions. Hering papers can now be secured from C. H. Stoelting and Co. The Milton-Bradley red and blue-green are nearly complementary.

Either the hand color mixer or the electric motor with adjustable resistance may be used.

Experiment No. 7. *Visual Adaptation and After-Images.*—The oblongs are Milton-Bradley colored papers pasted on 4 x 6-inch cardboard sheets. The small squares are made by pasting gummed cloth tape on the back of colored papers, trimming the edges, and then cutting the strips into ⅝-inch squares. Ink dots mark the center of both squares and oblongs. Materials are kept in sets in strong manila envelopes. On the envelope the make-up of each set is designated and the student is instructed to check his set for completeness when received and when returned.

Experiment No. 8. *The Visual Perception of Solidity and Distances.*—For the demonstration of double images we use 6-inch wooden rods (doweling) inserted in blocks 3 x 3 x 7/8 inches, and painted black. The background screen is a piece of the black cardboard used as a screen in mirror-drawing. It is supported by leaning it against the wooden rods described above.

The occlusion screen is a piece of black cardboard about 12 x 30 cm. with a piece of 3½ x 12 cm. cut from the middle of the lower edge. This edge is tacked to a wooden strip, so as to leave a slit 1 x 12 cm. above the strip. The pyramids can easily be made by bending and soldering the square "bus wire" commonly used for connections in radio sets. The base is made 5 x 5 cm., the height 5 cm., and the top 2 x 2 cm. The stereoscope and slides can be obtained from the C. H. Stoelting Co.

Experiment No. 9. *The Müller-Lyer Illusion.*—The Spencer Lens Co., Buffalo, N. Y., very kindly furnished us gratis with a number of 6-inch, 152+ mm. celluloid scales which are exceedingly convenient for measuring the length of the variable side of the illusion.

The board is similar in form to that of E. W. Scripture (*Yale Stud.*, 4, 1896, 101-103) and used by C. H. Judd (*Laboratory Manual of Psychology*, 1907, 1-21), and others. We make the board itself of ¼-inch "Compo board" (board center with paper covering) cut about 34 x 18 cm. This board will not warp or crack with ordinary use. Along the longer edges we tack strips of double thick cardboard 3 cm. wide. Upon this we lay a full size frame or border made of ordinary black cardboard 4 cm. wide, with opening 10 x 26 cm., and nail securely with double-pointed tacks through the under

cardboard strips and into the Compo board backer. The slides are of ordinary black cardboard a scant 12 cm. wide. Black is chosen so that use will not result in smudging the cardboard. With reasonable care such slides will last at least five years without the necessity of renewal. The longer variable slide is made 24 cm. long, the invariable, 20 cm. The figure is laid off on the slides by scale and dividers, and drawn in with white ink or thin water color paint. We use 3 cm. obliques. Fig. 13 is drawn approximately to scale. Each board and each set of cards is given a numerical designation to permit identification.

Experiment No. 10. *Special Phenomena of Vision I.*—Lowell's Test Type, the illiterate E-test type, astigmatic chart, the Ishihara Tests for Colour-Blindness, the Holmgren Worsteds and the perimeter may be secured from the C. H. Stoelting Co. Stillings *pseudo-isochromatische Tafeln zur Prüfung des Farbensinnes,* 17th edition, 1926, may be secured from Georg Thieme, Verlag, Leipzig.

Experiment No. 11. *Special Phenomena of Vision II.*—The Brewster stereoscope and Titchener slides may be obtained from the C. H. Stoelting Co. 3 x 5-inch index cards are the right size for investigating the blind spot. On the long diameter of the card place a small cross at $\frac{3}{4}$ inch from the left edge, a large black dot $\frac{1}{16}$ inch in diameter at $2\frac{3}{4}$ inches from the left edge, and a black square $\frac{1}{4}$ inch in diameter at $3\frac{7}{8}$ inches from the left edge. For negative after images use 2 gray cardboards about 20 x 28 inches, one with a red cross at the center. The cross should be 6 inches in diameter with arms 2 inches wide. The cardboards for brightness contrast should be at least 12 inches square. It requires a black

and a white cardboard with a 2-inch square of gray paper at the center of each, and a black cardboard with a $1\frac{3}{4}$-inch square hole at its center.

The color contrast frame is made by taking a piece of cardboard 16 x 28 inches and pasting to it, side by side, 4 pieces of colored paper (red, green, blue, yellow) 6 x 12 inches. These colored papers are centered on the cardboard and have a $\frac{1}{4}$-inch space between succeeding colors. Lengthwise of the cardboard and over the center of the colors is pasted a strip of gray paper $\frac{3}{4}$ inch wide. In another piece of very thick cardboard or plasterboard 16 x 28 inches, make 4 windows $5\frac{1}{2}$ x $11\frac{1}{2}$ inches separated from each other by sections of the cardboard 1 inch wide. Paste tissue paper over the windows. Then connect the two large cardboards together with hinges made of adhesive tape placed along the upper long edge. The cardboards should be so connected that when closed together the tissue paper will lie behind the windows of the first cardboard and flat against the colors. See page 76 of Titchener's *Text Book of Psychology* for a diagram of a contrast frame.

Experiment No. 12. *Imaginal (Ideational) Types.*—The pictures for description from memory may be made by pasting about 10 cut-outs of the largest pictures available on a single large sheet of cardboard.

Suggestions for materials are given in *Notes for Instructors.*

Experiment No. 13. *The Learning Curve.*—The materials are described in *Notes for Instructors.*

Experiment No. 14. *Habit Formation in the Maze.*—The designs for Maze "A" and Maze "B" are given

in *Notes for Instructors.* Two mazes have been used in order that both partners may serve as subjects if desired.

The mazes are made by sawing slots 1/4 inch wide in pieces of electricians' fiber, 12 x 9 x 1/4 inch. This material is light and does not readily break. Warping can be avoided by storage in a letter press. Each maze is labeled "A" or "B." Circular areas are sawed out and labeled "Start" and "Goal." Blind alleys are outlined in black to facilitate the counting of errors.

A tough manila double folder, one side of which goes below and the other above the maze, is used as a combined cover and tracing ground. The cover bears a warning that S is to be blindfolded before the maze is uncovered.

An ordinary wooden meat skewer is used as a stylus.

Experiment No. 15. *Mirror-Drawing.*—The mirror should be of good size, 6 x 9 inches or larger. The oblong base standard, already described for the knee jerk register and vernier pendulum, makes a good support. The long edge of the mirror rests on the base with its back against the upright rod, and the upper edge clamped to that rod with an ordinary right angle clamp. This gives the mirror the slight tilt required.

The screen is a piece of black cardboard 7 x 12 inches or thereabouts. It is gripped at one end by a test tube holder, and this holder is supported at the proper height on the rod of a second standard.

The star patterns may be mimeographed, two on each side of a large sheet. For use this sheet is folded at its center so that only one pattern shows. Arrows indicate the starting point and the direction for the movement in drawing.

Six-pointed stars form the pattern. The larger, which can just be inscribed in a circle whose diameter is 12½ cm., forms the outer edge of the path. The path is 2 mm. wide.

Experiments Nos. 16 and 17. *Methods of Investigating Memory.*—The materials used are described in *Notes for Instructors.*

Experiment No. 18. *Rational Learning.*—The materials are described in *Notes for Instructors.*

Experiment No. 19. *Animal Learning.*—The vertical maze * used in this experiment is pictured in Fig. 19. The main walls of the maze are built of ¾-inch boards and the back wall is ¼-inch thick. The front of the maze is covered with black window screen (H). Blind alleys are marked off by white bars (F) painted on the screen. The ascending runways (B) are constructed from sheet metal and the trap-doors (C) are made of heavy wire screen with ¼-inch mesh. On each trap-door is soldered a weight. These doors are lowered by means of slender wires (D). The food dish is placed on the upper level (just under letter H in figure). Door and doorway where the animal enters maze are shown by A′ and A. Door G is opened to remove animal at the end of a trial. The whole inside of the maze is painted black. Outside dimensions of the maze are 34½ x 4½ x 20 inches.

Experiment No. 20. *Learning and Interference in Card Sorting.*—The sorting placards are made of white

* This type of maze was built by Dr. A. G. Bills at the University of Minnesota in 1927.

cardboard 22 x 22 inches, quartered by ¼-inch wide black lines. (Dennison's binding tape can be used.) Each outer corner of a placard is indexed by the symbol of one suit of cards. Jacks, Queens, Kings and Jokers should be removed from the cards used in this experiment.

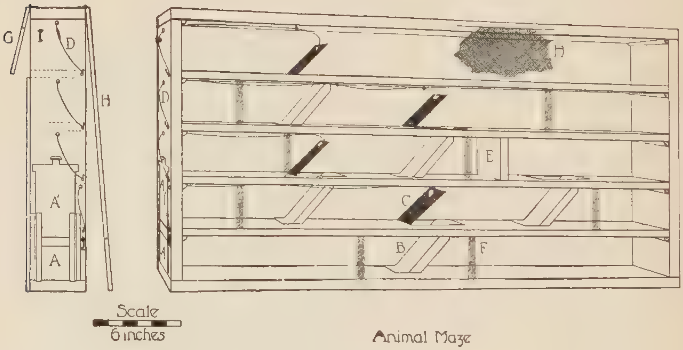
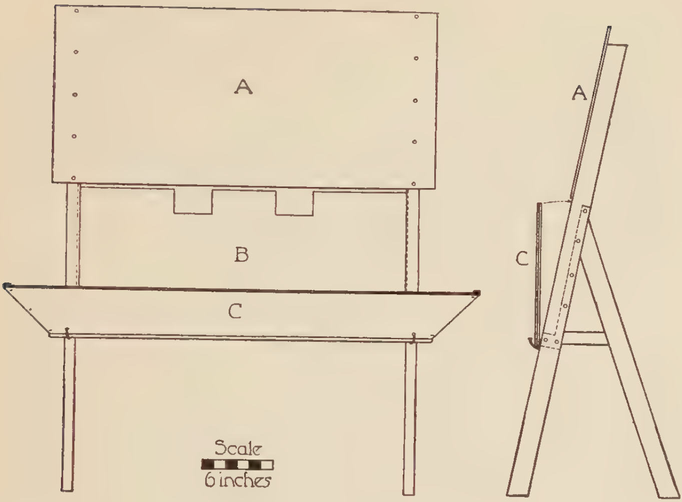


Fig. 19.

Experiment No. 21. *The Span of Visual Apprehension.*—The stimulus materials are described in *Notes for Instructors.* The exposure frame for presenting stimuli is pictured in Fig. 20. About 10 stimulus cards are placed in compartment B. The top card is removed after each exposure. To expose stimuli, E stands behind frame, grasps door C at the two ends and turns it forward and downward. E can easily watch a timer while the exposure is made. The frame is 40 inches high and 32 inches wide. A and C are heavy gray cardboards. B has 2 notches cut in the top to facilitate the removal of stimulus cards. Cards in B are prevented from falling forward by thin strips of metal projecting inward from the legs at either end of the frame.



Exposure Frame

Fig. 20.

Experiment No. 22. *A Group Intelligence Examination.*—Information with regard to obtaining blanks and with regard to instructions and results for the test is given in *Notes for Instructors.*

Experiment No. 23. *The Measurement of Musical Talent.*—The Seashore records may be obtained from the Columbia Graphophone Co. The musical self-rating scale is described in *Notes for Instructors.*

Experiment No. 24. *The Measurement of Reading Ability.*—Information with regard to obtaining test blanks is given in *Notes for Instructors.* Exercises for the study of eye movements in reading are given below. (These paragraphs are taken from Science, 1926, N. S., Vol. 63.)

PRACTICE EXERCISE

The gaseous products set afloat by the family furnace are likely to have a corrosive action on the book bindings in the family library. According to experiments conducted at the Bureau of Chemistry, the products of combustion which pollute the air in large cities have a very deteriorating effect on the leather of bindings. This may be counteracted by applying dressings, either while the leather is being made or when the book is bound.

EXERCISE A

The night-blooming cereus, one of the most beautiful flowers of the desert, is not wholly a flower of the night, as has commonly been thought. The annual exhibit of scientific work at the Carnegie Institution of Washington includes two photographs of these rare plants being visited by bees that fly only in the daytime. Dr. D. T. MacDougal, of the institution, explains that the flower does not close with the coming of the first signs of dawn, but remains open for a time, usually giving the bees an hour or two of early daylight to work by, and thus freeing itself of total dependence upon the night-flying moths for the highly important process of pollination.

EXERCISE B

A cook stove built by the National Geographic Expedition to Mt. Katmai five years ago was found and used by an explorer last fall. It was as hot as ever and cooked food perfectly, without fuel or smoke. The stove is a two-foot hole dug out of a boiling fumarole of steam in the valley where over a thousand of the steaming potholes are located. The discovery of this cook hole and its identifying mark proves that the volcanic regions about Mt. Katmai are not cooling to a perceptible degree.

Experiment No. 25. *The Attention Value of Advertisments.*—Suggestions for the selection of advertisements are made in *Notes for Instructors.*

Experiment No. 26. *The Affective Value of Colors.*—The colors are pasted on lightweight white or gray cardboard 8½ x 11 inches. The Milton Bradley or similar colored papers are used. For single colors use colored papers 4 x 5 inches. For pairs of colors, have each color 2½ x 4 inches and paste the two colors side by side. Number each stimulus card near the top with figures about 1½ inches high. The exposure frame employed in experiment No. 21 is used for exposing the colored stimuli.

Experiment No. 27. *The Affective Value of Musical Intervals.*—The results quoted were obtained for intervals given with the Ellis Harmonical, untempered tuning. It is described in H. Helmholtz, *Sensations of Tone* (translated by A. J. Ellis), 1895, p. 17. A tonometer, organ, or piano with tempered tuning may be used instead of the harmonical. The musical intervals and the rating scale to be used in this experiment are described in *Notes for Instructors.*

Experiment No. 28. *Free Association.*—The materials are described in *Notes for Instructors.*

Experiment No. 29. *Associative Responses as Diagnostic.*—The materials are described in *Notes for Instructors.*

Experiment No. 30. *Judgment of Intelligence from Photographs.*—The photographs to be used are given

below, pp. 386-387. The list of I.Q.'s is given in *Notes for Instructors.*

The Pintner series of photographs was dropped and a new series prepared for this experiment in order to control to better advantage the factors of sex, age, and such secondary criteria as costume, size of face, etc. The reviser gratefully acknowledges the coöperation of the Institute of Child Welfare, at the University of Minnesota, in the securing of the pictures and the test scores.

Experiment No. 31. *Judgment of Emotion from Photographs.*—The photographs are published in *Psychological Review,* Vol. 21, 1914, pp. 33-41. Copies may be secured from C. H. Stoelting and Co. The emotions intended to be represented are listed in *Notes for Instructors.*

Experiment No. 32. *The Measurement of Art Talent.*—Test books for the Meier-Seashore Art Judgment Test, record sheets and full instructions (containing statements of difference between pictures and suggestions for interpretation of results) for giving may be obtained at cost from the Bureau of Educational Research, University of Iowa, Iowa City, Iowa.

INDEX OF SUBJECTS

INDEX OF NAMES